

Lessons from Huntingdon

The virtual destruction of an animal-testing company by activists and terrorists has highlighted again the power of fundamentalist minorities. Industry and government have failed to respond adequately to the public challenge.

A company that sets itself up in Britain, one of the most animal-protective nations in the world, to expose tens of thousands of animals every year to tests that cause pain or death might be said to be asking for trouble. But such activities are not only legal, they are also an unpleasant necessity.

If, however, such a company is short-sighted enough to conduct its tests in a manner that is both cruel and illegal, it fully deserves to lose the backing of customers, investors and the public. The pressures on the animal-testing contractor Huntingdon Life Sciences (HLS) from animal-rights activists appear to have been stimulated by a succession of undercover investigations during the past 10 years which revealed, most powerfully in a television broadcast in 1997, that the firm had much to answer for.

Talk to some of its customers today, and you find their faith in its practices has been restored by its changed management. However, the number of its customers has diminished, thanks to a relentless, vicious and clever campaign that followed on the heels of a previous campaign which had succeeded in closing down a supplier of cats used in biomedical research. A look at the website of the group Stop Huntingdon Animal Cruelty (SHAC) (<http://www.freespeech.org/sharelist/SHACTEMP/>) reveals the continuity of the campaign, while a read of its "demo reports" gives a taste of the attitudes behind the demonstrations: "The timing was perfect. As the staff from HLS were leaving work on Thursday evening, over thirty animal rights activists besieged the site. They surrounded a car that was half way out of the gateway. The three police officers who were on the scene could do nothing but stand by and watch as the car was forced back into HLS and security guards struggled to close the gate. ... It took over 45 minutes before enough police reinforcements arrived to enable the workers to leave in relative safety. However, a separate group of activists were waiting further down the road and ambushed some of the workers. Several of the workers had their car windows smashed. ... Later on in the evening people paid a visit to [name deleted], head animal technician at HLS' Occold lab to remind him to have a filthy Christmas and to advise him to leave HLS. We think he got the message!"

Illegal tactics

The tactics described have ranged from breaking into the factories of HLS's suppliers, distributing leaflets to the neighbours of HLS workers, and threatening HLS investors and bankers in the United Kingdom and North America. Illegal tactics by rogue activists, disowned by SHAC, are described on its website by means of quotes from news stories — for example, an HLS executive being temporarily blinded by an unidentified substance thrown at him.

To seek a rational debate with those behind the protests is to miss an essential point. Although the anti-HLS campaign appears to be well focused and have clear originators, it has attracted some, again disowned by SHAC, who belong to what is more a movement than an organization, referred to as the Animal Liberation Front, which began in 1976 — some of its self-nominated fundamentalist 'members' believe that animals are more deserving of protection than people, and that "to drink milk is to have blood on your lips".

A good portrait of the Animal Liberation Front's 'members' can be found at <http://www.guardianunlimited.co.uk/Archive/Article/0,4273,3824140,00.html>

Side by side with these terrorists are people who have far less extreme views, but who are appalled by Huntingdon's past reported misdemeanours. Whatever the type of anti-HLS campaigner, their style in television debates has been well calculated: avoid specifics, repeat propaganda statements as a mantra and thereby trade on the gut feelings of the public against animal testing.

Government intervention

The UK government has belatedly woken up to the effectiveness of such tactics, which have brought HLS to the brink of extinction. Pressure from the government has helped HLS find last-minute, albeit anonymous, financial help. History suggests that, had it not been for the fact that the company's demise would have sent a damaging signal about Britain as a place in which to do biology-based business, the British government would not have lifted a finger.

The campaigners would have moved on to another target had HLS folded, but as it is, the activists have sworn to keep the pressure on the company's customers and to track down its new anonymous US investors. As such actions continue, the mass media, having appropriately broadcast the past shortcomings of HLS, need to ensure that the true character of these campaigns and the people behind them are also kept in the public eye.

Furthermore, the pharmaceutical industry needs to rethink its public tactics — unless it intends to concentrate future efforts in the many countries where the regulation of animal research is more lax than in the United Kingdom. But assuming that the industry still wishes to work in Britain, it should adopt a higher profile, starting with the Internet. The web is full of information about animal cruelty, real and alleged, but informed responses to specific campaigns are sparse. Some organizations — not least the Research Defence Society — offer general facts and figures. But it is hard to find any statement from industry — the websites of the big companies and of the Association of the British Pharmaceutical Industry appear to provide no information that might help the defenders of animal testing in their cause. The SHAC website includes letters to SHAC from Merck Pharmaceuticals UK and from Servier UK stating that they are not contracting HLS to test their products. If terrorist tactics were responsible for their decisions, the companies should work with the government to obtain support and reverse them.

Finally, the research community and its regulators, in and outside Britain, should ask themselves what the lessons of these events are for the conduct of animal testing and research. There is a need to maintain a higher profile about the continuing progress in the refinement, reduction and replacement of the use of animals. Researchers dependent on animals — together with their funding agencies — should also work to overcome the inertia of experience and expertise, and aim to keep animal testing and models as far down the animal kingdom's hierarchy of organismal complexity and cognitive ability as their scientific goals allow. ■





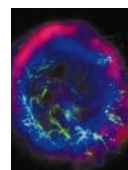
Hot seat
Animal researchers
seek government help
against protesters
p443



About face?
Will Bush's health
secretary still back
work on stem cells?
p444



Physical fitness
CERN survey finds
particle physicists in
great demand
p445



News feature:
nuclear physics
Cooking up elements
in supernovae
p449

Agencies face uphill battle to keep United States free of BSE

Meredith Wadman, Washington

US government agencies are confronting a tough test of their much-vaunted ability to maintain a safe food supply, as they scramble to block possible paths for the entry of mad cow disease into the United States. Their latest action is to extend a ban on blood donors from European countries.

Until late last year the disease, more properly known as bovine spongiform encephalopathy (BSE), was thought to be largely confined to the United Kingdom, Switzerland and Ireland. In the United Kingdom, more than 80 people have died from variant Creutzfeldt-Jakob disease (vCJD), an analogous human disease presumed to be contracted by eating infected cattle.

But the BSE panic is fanning out across Europe, with Italy the latest country to succumb (see below). Its implications are registering with the American public for the first time, putting intense pressure on agencies charged with protecting food safety. These, primarily the Food and Drug Administration (FDA) and the United States Department of Agriculture (USDA), confront a challenge that is logistically complex, politically delicate and scientifically uncertain.

"The FDA and USDA face a tough balancing act against a frightening, little-known enemy. It's as tough as any [challenge] in the last decade," says Arthur Caplan, director of the University of Pennsylvania's Center for Bioethics. He chairs a committee that advises the Department of Health and Human Services, the FDA's parent agency, on blood safety.

An FDA advisory committee last week recommended extending a ban on blood donors to include people who have, since 1980, spent ten years in France, Portugal or the Republic of Ireland. Previously, the 1999 ban had applied only to those who had spent six months or more in the United Kingdom between 1980 and 1996.

It is not known whether the disease can be spread through blood transfusion, although the epidemiological evidence so far makes this unlikely. But, says Paul Brown, chair of the advisory committee and director of the Laboratory of Central Nervous System Studies at the National Institute of Neurological



Cowed: BSE could unseat the US cattle industry.

Disorders and Stroke, "the committee chose a compromise conservative position. They are assuming there's a significant potential risk of exposure in Europe."

The committee had been pressed unsuccessfully by the US Red Cross and others to bar donors who had spent less than six months in the United Kingdom and to extend the ban to the whole of Europe.

BSE and vCJD are caused by abnormally shaped brain proteins or prions. These cause neurological deterioration, leading to dementia, uncontrolled movement and

rapid death. There is no diagnostic test for the diseases, which are confirmed by microscopic examination of the brain after death.

No cases of BSE or vCJD have been diagnosed in the United States, according to the USDA and the Centers for Disease Control and Prevention. The USDA has examined the brains of 11,900 cattle in the past decade.

"Currently we are free of mad cow disease in cows and variant CJD in humans. Every possible effort must be expended to keep us in that category," says Sidney Wolfe, who heads the consumer watchdog Public Citizen Health Research Group in Washington DC.

Regulation of the many routes by which BSE might enter the United States is complex. The FDA monitors animal feed, blood products, vaccines and dietary supplements. The USDA is responsible for safety of meat and the health of 900 million US cattle.

"If we get BSE in this country, our cattle industry goes down the tubes," says Jim Rogers, a spokesman for the USDA's Animal and Plant Health Inspection Service.

The FDA has no power to stop the manufacturers of dietary supplements using processed cow parts, including brain, in

Italians drop beef as first cow tests positive

Sergio Pisto, Turin

Italy's demand for beef has collapsed dramatically following last week's discovery of the country's first BSE case. The finding came two weeks after the government introduced tests for all cattle over 30 months old.

Butchers reported a much sharper fall in demand than the estimated 30% decrease that followed the BSE crisis in France last October (see *Nature* 408, 392; 2000). One association of butchers said that consumption fell by 70-90% in the days after the discovery.

The infected animal was a six-

year-old cow from a farm near Brescia in northern Italy. It had shown no symptoms of BSE but was found to be positive in a prionics test performed on brain tissue. The initial result was confirmed by researchers at Italy's national BSE reference centre in Turin.

To ease the pressure on laboratories that are now running as many as 1,500 tests a day, the government is encouraging farmers to slaughter older animals without testing. Only ten veterinary centres in the country are equipped for rapid testing, but Umberto Veronesi, the health

minister, said that 13 more will be established in the next few weeks.

The government is also working to computerize a national bovine register, which currently covers about four-fifths of the cattle herd, to track each animal's origin and history. The testing age is also to be lowered to 24 months, in accordance with a European Union directive.

According to Veronesi, only one of 6,000 tests performed by the end of last week was found positive, indicating that the rate of infection is low. But he said 50,000 tests were needed to get a reliable estimate of the rate. ■

their products. Similarly, a USDA import ban on bovine tissues applies only if they are destined for use in food and medical products, not if they are to become dietary supplements. One nationally distributed supplement lists 17 bovine organs, including brain, as ingredients.

The United States has not imported British beef since 1985, and it banned imports of live cows from BSE-affected countries in 1989. But a recent report by a scientific advisory panel to the European Union says that tonnes of high-risk European beef products are shipped to the United States annually. They include gelatin, collagen, semen and albumin, and are used in products including cosmetics and vaccines. (In December, the FDA forbade US drug-makers to use bovine serum from BSE-affected countries in certain vaccines, but vaccines produced before the ban are still in use.)

Besides the loopholes for exports, measures to protect the US herd are sometimes disregarded. BSE is believed to have spread through the practice, now prohibited, of using cow and sheep parts in cattle feed. The FDA announced on 10 January that hundreds of US feed producers are not complying with 1997 regulations enforcing the ban.

Beyond the challenge of enforcing current regulations, the US agencies need to examine how to plug loopholes, and to determine the best way to spend limited resources fighting a dimly understood enemy. How much effort should be spent in developing a screening test for the diseases or finding a treatment, or on blocking any imported or home-grown vehicle that might spread the diseases?

Deer and elk herds in the western United States, for instance, are being afflicted in growing numbers with an analogous ailment called chronic wasting disease. Last week, the committee that extended the blood-donor ban was asked to consider if and how blood donations from people exposed to these deer and elk should be restricted. It decided there were not yet enough data to implement restrictions on them.

All of these issues "get into foreign policy questions, trade issues, regulation of hunters, which Americans hate. These are all politically tricky," says Caplan.

At the advisory committee meeting last week, scientists complained of the lack of data on which to base their decisions — including the future curve of the epidemic in Europe, and the risk of exposure in other European countries compared with the United Kingdom. "These decisions ... are really guesses," said Stanley Prusiner, who won a Nobel prize in 1997 for discovering prions and elucidating how they cause disease.

Public have right to see data from all gene trials, says FDA

Paul Smaglik, Washington

The US Food and Drug Administration (FDA) last week put forward proposals to make clinical trials in gene therapy and xenotransplantation more transparent. The move drew support from some academic scientists, but criticism from the biotechnology industry.

The proposal — which interested parties have 90 days to comment on — would close a loophole that allows some data from privately funded trials to be concealed. Comments can be sent to fdadockets@oc.fda.gov.

Public disclosure became an issue in 1999, after a teenager died in a clinical trial at the University of Pennsylvania.

Adverse effects from federally funded gene-therapy trials are meant to be reported to the National Institutes of Health (NIH) and the FDA. But after the death, NIH investigators learnt that hundreds of such events had not been reported to the agency, which makes data public through the Recombinant DNA Advisory Committee (RAC). Most had only been reported to the FDA, which keeps

them confidential. Researchers complained that the agencies had different reporting standards. Last week's announcement aims to harmonize those requirements, and expand them to xenotransplantation.

The new policy would ensure that data on adverse events from all gene-therapy trials would be made public. Some are already made public through an NIH database: the change would ensure that this database contains information from all ongoing trials.

Inder Verma, president of the American Society of Gene Therapy, applauds the FDA's proposal, describing it as necessary to regain public trust. "Anything that will make it a more open process is good," says Verma, a Salk Institute gene therapy researcher.

Michael Werner of the Biotechnology Industry Organization (BIO), which opposes the proposal, says it raises privacy concerns. The BIO favours releasing aggregated safety data and reports on trends in adverse events rather than information about individual trials.

♦ <http://www4.od.nih.gov/oba/clinicaltrial.htm>

Students fuzzy about the future

Mark Schrope

Students at US graduate schools are happy to be doing PhDs and love their research, a new survey says. But they do not get enough teacher training or career advice, and their rapport with staff has room for improvement.

Just 3% of those surveyed admitted to being unhappy with their decision to pursue a PhD and 93% said their research interested them "a great deal". But more than a third might consider changing universities, advisers or dissertation topics

if given the chance to start again.

The survey, carried out by the University of Wisconsin at Madison for the Pew Charitable Trusts, polled more than 4,000 students in the third year or above in 27 graduate programmes.

"We were trying to bring the experiences and voices of students into the debate," says lead author, educationalist Chris Golde.

Only about a third felt their programmes were preparing them to teach lecture courses, compared with three-quarters who felt up to the task of conducting research. Students in the 'hard sciences' were more likely than those in the social sciences or humanities to be required to teach, but were less likely to receive any training.

Under half had a clear understanding of what (or how long) it would take for them to graduate. Molecular biologists and chemists were the fuzziest on these points, probably reflecting the unpredictable nature of laboratory research, according to the report.

The report's authors call for better communication between staff and students, including annual reviews to discuss needs, expectations and progress — an opportunity now only available to about half the students. They also urge students to voice their needs and expectations.





Burning issue: some activists resorted to illegal means in expressing opposition to animal research.

Animal-rights protests raise calls for UK government move

Georgina Kenyon, London

The British government is under intense pressure to do more to protect people associated with companies that perform experiments on animals, following last week's near-collapse of the Cambridgeshire-based company Huntingdon Life Sciences (HLS).

HLS came close to bankruptcy when one of its main backers, the Royal Bank of Scotland, withdrew its loan support. The bank had faced a campaign by animal-rights protesters that had targeted its directors.

The fate of HLS, one of the United Kingdom's largest animal-research contractors, is being seen as a critical test of whether animal research can be protected from an array of protest actions. These have ranged from legal demonstrations, backed by legitimate animal-rights groups, to letter-bomb campaigns by anonymous extremists.

Jack Straw, the British Home Secretary, announced on 17 January that the government was considering tougher laws to stop two tactics being used by animal-rights activists: protests outside the homes of company directors and threats sent to them by mail.

But the Research Defence Society, a group representing scientists involved in animal experimentation, attacked the government's proposals as "too little and too late".

Barbara Davies, the society's deputy director, also questioned whether the government would actually implement changes in the law. "We need to see the government's words followed by action," she says. "We have been pressing for several months for legislation to outlaw [some] activists' violent campaigns against scientists and shareholders."

Groups opposed to animal research had recently extended their campaign from HLS itself to the bank. The bank had been considering whether to renew its loan facility of £22.5 million (US\$33 million) to HLS, which

expired last Friday. But HLS reportedly secured new funding from unidentified backers in the United States. Groups such as Stop Huntingdon Animal Cruelty, whose policy is to support only peaceful and legal protests against animal research, say they will pursue their campaign against the new financiers if they are identified.

HLS was criticized for animal mistreatment in a 1997 British television programme, which showed HLS employees throwing a beagle against a wall and seriously harming other laboratory animals. The company has since replaced its management, and its supporters say its standards of animal care have improved.

A spokesperson for the Association of the British Pharmaceutical Industry says: "Nothing justifies the way these protesters have acted. We have to test on animals at this point in time. And if Huntingdon is forced to close, considerable damage will be done to the future of UK medical research."

The association is currently in discussions with the government and other industry representatives over possible measures to prevent further demonstrations, and ease pressure on research laboratories to close.

The government has also been criticized by researchers for sending what they regard as ambiguous signals on its general policy towards the animal-rights movement. In particular, it failed to publicly support the concept of medical research on animals last year, when several pension funds that managed money for government employees withdrew money from HLS on ethical grounds.

Nancy Rothwell, a professor of physiology at the University of Manchester, says: "The government should introduce the possibility of withholding shareholders' and employers' names from the public. If HLS shuts down, the knock-on effect is a great concern for the research community."

Max Planck Society broke pledge to ask Jews back after war

Vera Bettenworth and Alison Abbott, Munich

Germany's leading research organization reneged on a 1948 promise to contact Jewish scientists expelled from its laboratories during the Third Reich and offer them their jobs back, according to new research.

In fact, the research suggests, the Max Planck Society (MPS) tried only to tempt star scientists back to its fold. Most of these, including Albert Einstein, refused to return to Germany. Einstein thought his invitation was a whitewash attempt, and said he had "an irresistible aversion to being part of German public life again".

Historian Michael Schüring has uncovered 85 cases of Jews or 'political undesirables' being expelled from the Kaiser Wilhelm Society, the MPS's predecessor. He adds that the true figure is probably higher, as documents may have been lost and witnesses are dying out. Only a small proportion of the cases he unearthed appear to have been contacted after the war, Schüring says.

Of these, only three agreed to return to work in Max Planck institutes, and one, geneticist Max Ufer, changed his mind when he learnt of the fascist past of some of the scientists he would have had to work with. Others were willing to have some level of scientific contact, for example as guest researchers or lecturers.

In the first three decades after the war, some dismissed members claimed compensation for damage to their careers, on their own initiative. Thirty-three such cases are documented in the state and MPS archives. Of the 23 claims in the MPS archives, 14 were accepted in the form of one-off payments or pensions. All payments were made on a voluntary basis, the MPS refusing to accept legal obligation.

The research was commissioned by MPS president Hubert Markl, who has said the society "must definitely ask whether everything was done in the early years of the Max Planck Society to compensate for the shameful behaviour of the mid-1930s".



Einstein reaches New York in 1933: he refused entreaties to return to Germany.

Health and energy chiefs sail through Senate

Matthew Davis and Irwin Goodwin, Washington

The two cabinet secretaries who will oversee most US science in the new Bush administration sought confirmation from the Senate last week.

But as Tommy Thompson, the health secretary, and Spencer Abraham, the energy secretary, sailed through their confirmation hearings, little was revealed of their plans for US research. Abraham was duly confirmed on Saturday, with Thompson expected to follow early this week.

Although opposed to abortion, as governor of Wisconsin Thompson was a forceful advocate for human embryonic stem-cell research. He publicly lauded a scientist in the state who pioneered a technique for extracting stem cells from embryos, and stood up to anti-abortion groups opposed to the work.

But as President Bush's new Secretary of Health and Human Services, one of Thompson's first tasks could be to stop efforts by the National Institutes of Health (NIH) to provide federal funding to explore the therapeutic potential of stem cells. His department includes the NIH, and he will play a key role in Bush's selection of its next director.

"It's hard for me to see how he can do a total about face and condemn this research as immoral when he is on record as having applauded this advance," says Paul Berg, a molecular biologist at Stanford University who heads the public policy committee at the American Society for Cell Biology.

But with no one seeking to upset his popular nomination, none of the senators asked Thompson for his views on statements from the Bush presidential campaign that the new administration would end the NIH's support of stem-cell research. And asked afterwards by a reporter whether he would be as supportive of stem-cell research



Hands off: health secretary Tommy Thompson refuses to be drawn on stem-cell research.

in Washington as he had been in Wisconsin, he refused to answer.

Proponents of stem-cell research were relieved the contentious issue was kept out of the limelight. They feared that putting Thompson on the spot would have forced him publicly to endorse the Bush campaign position. But the powerful anti-abortion lobby believes that Bush will still end the NIH initiative.

Ironically, the NIH's plan to allow stem-cell research would depend on James Thomson of the University of Wisconsin at Madison to provide cell lines. Thompson described Thomson in his 1999 State of the State address as one of the "bold pioneers who are leading Wisconsin into the next millennium".

Sources in Wisconsin say that Thompson is quite knowledgeable about stem-cell research and could play a pivotal role in any internal administration debate.

The nomination of former senator Spencer Abraham as energy secretary raised some eyebrows because he had twice sponsored legislation to abolish the department

he will now lead.

But Senator Jeff Bingaman (Democrat, New Mexico), who chaired the confirmation hearing, said that Abraham "has since seen the light and come to understand how much the department does for our energy security, our national security, our economy and our scientific and technological prowess".

Nuclear weapons labs get one-third of the department's \$19 billion budget, and the Office of Science, which supports non-weapons research, spends another \$3 billion of it. Last year, Abraham sponsored a bill to raise Department of Energy funds for science.

"The science and technology programmes at the department have been widely praised and justly so," Abraham told the hearing.

Concern has been expressed about Abraham's lack of any obvious qualification for the job. But Michael Lubell, a physicist at City University of New York and lobbyist at the American Physical Society, says Abraham is a "political pragmatist" who will probably take advice on science from Michigan congressman Vernon Ehlers, a former physicist. ■

Canadian minister keeps researchers guessing

David Spurgeon, Montreal

Canadian scientists are watching closely to see whether the new industry minister in the Liberal government, which was re-elected last November, will continue with the pro-science policies of recent years.

In the new government, Brian Tobin, the politically ambitious former premier of Newfoundland, replaces John Manley, who as industry minister was in charge of the nation's science and technology policy.

Under Manley's influence, science budgets increased and national networks such as the Canada Foundation for Innovation and the Canadian Institutes for Health Research were set up (see *Nature* 405, 722; 2000), raising

morale among researchers.

But Tobin, who is perhaps best known for encouraging the Canadian coastguard to fire on Spanish fishing boats in a 1995 dispute known as the 'turbot war', is seen as an aggressive politician and an unknown quantity on research issues.

John de la Mothe, a science-policy expert at the University of Ottawa, is concerned that Tobin's confrontational approach might damage some of Manley's achievements, citing a "difference of style between the two".

But Ron Freedman, an official at The Impact Group, a consultancy firm specializing in science and technology, says: "Tobin was premier of Newfoundland, and a large part of their development strategy was



Hard to gauge: industry minister Brian Tobin.

based around the new economy companies." Freedman adds that the new minister is "probably well aware of the issues" and will "pick up very quickly" on the government's existing research strategy.

John Reid, president of the Canadian Advanced Technology Association, says: "Tobin's made it quite

clear that he respects what Manley has done and eventually is going to carry that forward." ■

Assessment ups the ante on climate change

Quirin Schiermeier, Munich

Global warming is liable to become an even more acute problem than anticipated, according to the new assessment of the Intergovernmental Panel on Climate Change (IPCC).

Whereas global average-surface temperature has increased by 0.6 °C during the past century, it is set to rise by between 1.4 °C and 5.8 °C by 2100, according to the report. There is "new and stronger evidence" that global warming is caused by human activities, it adds.

"The projected rate of warming is much larger than the observed changes during the twentieth century and is very likely to be without precedent during at least the last 10,000 years, based on palaeoclimate data," it says.

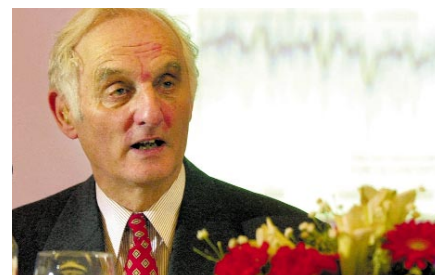
The report was finalized and unanimously accepted last week in Shanghai by 150 scientists and government representatives of the IPCC's working group on the science of climate change. The working group's third assessment report (the first two were published in 1990 and 1995) was three years in the making, involving 123 lead authors from around the world, and more than 500 contributing authors.

The new estimation of the probable range of global warming differs considerably from that in the 1995 report, which had projected a maximum temperature rise of 3.5 °C. The new projection — almost the same as the figures leaked to the press last October — is based on advances in climate modelling and on data that have become available from empirical work since 1995.

But the authors of the report remained cautious about detailed predictions of possible climate instabilities on a more regional scale. Frequency and intensity of extreme local weather events, such as hurricanes and heavy rainfall, have been excluded from the projection, owing to "scientific uncertainties". New research on possible discontinuities in the ocean circulation, such as a shut-off of the thermohaline circulation of the North Atlantic (see *Nature* 409, 153–158; 2001), was also left unconsidered.

The report arrives after international talks last November failed to reach agreement on implementing the 1997 Kyoto Protocol to reduce greenhouse-gas emissions.

John Houghton, co-chair of the IPCC working group on the scientific assessment



Stern warning: John Houghton unveils the revised predictions on global warming.

of climate change, is happy that the "serious news" of the report was accepted by the representatives of governments, including the United States and Saudi Arabia, which, in the past, had been reluctant to accept evidence for man-made climate change.

"It is now widely undisputed that the increased intensity of the hydrological cycle will lead to a rise in sea levels, and, in some regions, to more frequent floods and droughts," Houghton says. "I hope that this insight will do a great deal to convince governments that actions further to the Kyoto Protocol are urgently needed." ■

Demand for physicists on the rise

Alison Abbott, Munich

What becomes of research students who study the fundamental forces of the Universe? According to a survey of students who pass through CERN, the European Laboratory for Particle Physics, esotericism proves no bar to lucrative employment.

Half of the 700 students surveyed landed well-paid jobs in high-tech industries, and the other half stayed employed in academic research. None were left looking for work.

The survey confirms other indications that physicists are in demand in the private sector, says Alex Bradshaw, vice-president of the German Physical Society. The society has been warning for several years that there will soon be a shortage of physics graduates. Fear of unemployment is believed to be one of the reasons fewer students have enrolled in the physical sciences in Europe during the past decade.

The CERN survey tracked students from 19 countries who worked on Delphi, one of four major particle detectors on CERN's Large Electron-Positron collider. Whereas, in 1996, 4.8% of students were without a job immediately after graduation, zero unemployment was recorded in 2000. The percentage of those employed in the private sector increased from 43% to about 50% during these five years.

"High-tech industries want to acquire the practical skills that students learn at CERN," says Tiziano Camposi, spokesman for the Delphi experiment.

The survey also showed that the number of female students on the experiment increased over the years to 20%, and that they followed the same early career paths, with the same levels of responsibility, as the male students.

Bradshaw hopes that the experience of students at CERN will encourage more interest in physics among school-leavers. An upturn appears to have already started, he says, with a significantly increased enrolment in university physics programmes in Germany this year. ■



Glowing future: it may be esoteric, but particle physics provides strong career opportunities.

Stem-cell research to start in Britain

David Dickson, London

British scientists have been given the go-ahead to carry out research on cloned human embryos aimed at the treatment of diseases. Such research — which is likely to remain outlawed in most of the rest of Europe — could now begin in Britain before the end of the year.

On Monday, the House of Lords voted to approve the government's proposal that such research should be allowed under the Human Fertilization and Embryology Act. A similar vote was passed last month by the House of Commons (see *Nature* 409, 5; 2001), and scientists will be able to apply for licences for the research shortly.

The margin of the vote in the Lords — where the proposed amendment was approved by 212 votes to 92 — surprised observers who had anticipated a closer vote, particularly in the light of calls for caution from religious leaders, including the Archbishop of Canterbury, the head of the Church of England.

Some critics appear to have been mollified by a Lords' plan to set up a committee to review the ethical issues that the research raises. The government has also promised to introduce "as soon as possible" a bill to ban reproductive cloning. ■

Partnership seeks to build supercomputer for proteomics

Washington Sandia National Laboratory in New Mexico, Compaq Computer Corporation and Celera Genomics last week joined forces to build a supercomputer tailored to proteomics.

The Department of Energy will provide \$10 million over four years to build a machine that can perform 100 trillion operations per second. Project leaders hope that, with some technological breakthroughs, the machine could eventually run ten times faster than that.

Although comparisons to another supercomputer currently in development — IBM's Blue Gene (see *Nature* 402, 705; 1999) — seem inevitable, they may not be completely appropriate. Each machine's performance is measured differently, and the two computers have been designed for different tasks. Blue Gene may ultimately be the faster computer, but the newly proposed machine will probably have more memory.

Speed is more important for Blue Gene because it is designed for problems that require many simultaneous calculations, such as the simulation of protein folding. Memory is more important for the new

machine, its designers say, as it will be used to process huge volumes of protein and genome data.

Rule change eases listing on London stock exchange

London Britain's financial regulators have proposed easing the rules under which small biotechnology and other science-based companies can be listed on the London Stock Exchange. At present, a company must first achieve one of four formal milestones — such as having two products at an advanced stage of development.

The Financial Services Authority now says this requirement should be abandoned, reflecting a trend away from product development to, for example, "the operation of genomic sequence databases to assist gene therapy development". A company would merely need to make a detailed disclosure of milestones "that are appropriate to their particular area of scientific research activity".

♦ <http://www.fsa.gov.uk/pubs/cp/81>

Australian to head optical observatory

Washington Australian astronomer Jeremy Mould has been selected as the next director of the National Optical Astronomy



New observer: Mould is heading for US.

Observatories (NOAO) in the United States.

Mould, director of the Australian National University's Mount Stromlo and Siding Springs Observatories, is to succeed Sidney Wolff at the NOAO's headquarters at Tucson, Arizona.

The appointment was made by the Association of Universities for Research in Astronomy,

which operates the observatory for the National Science Foundation (NSF). The NSF is expected to confirm the selection shortly.

Colorado gets \$250m gift to fight cognitive disabilities

San Diego In one of the largest ever gifts to a public university in the United States, the University of Colorado has received \$250 million over five years for research and engineering for technology to help people with cognitive disabilities.

The gift was made last week by William Coleman III and his wife, Claudia, who are

both executives with high-tech firms in Silicon Valley. He is chairman and founder of BEA Systems Inc., an Internet software firm, and was a founder of Sun Microsystems Inc.

The money will initially be used by scientists and engineers at the university's main campus in Boulder and at the health science centre in Denver.

Oil slick threatens Galapagos coastline

London Ecuador is appealing for international assistance to help control an oil spill just 800 metres off the coast of the Galapagos Islands. The oil is threatening the archipelago's unique coastal ecosystem.

The spill occurred on 16 January when the Ecuadorian tanker *Jessica* struck rocks off the island of San Cristobal. So far, 144,000 gallons of diesel and bunker — a type of heavy fuel — have spilled into the sea, and currents are carrying the slick south towards the archipelago's largest sea-lion colony.

Volunteers from the Ecuadorian National Park Service are preparing to treat oil-coated animals and evacuate them from the island if necessary. The spill is also threatening marine iguanas, which are unique to Galapagos, as well as various birds including blue-footed boobies and pelicans.

Spread money round more, says Royal Society

London The Royal Society warned this week that the diversity and strength of British science could suffer if government agencies continue to concentrate their resources in a small handful of elite research universities.

In a report, the society urges that universities which get middling grades for research excellence should not be excluded from receiving money from the Higher Education Funding Council for England.

An increasing share of research money has been going to Oxford, Cambridge and some London colleges, notes the report. This leaves less for middle-ranking institutions which, according to John Enderby, vice-president of the society, "make an important contribution to the dynamism and diversity that is necessary to maintain the vitality of UK university research".

Energy department dives into Biosphere 2

San Diego The US Department of Energy last week signed an agreement with Columbia University in New York for a \$1 million project to assess whether the Biosphere 2 Center in Arizona can be developed as a national laboratory for Earth systems research.



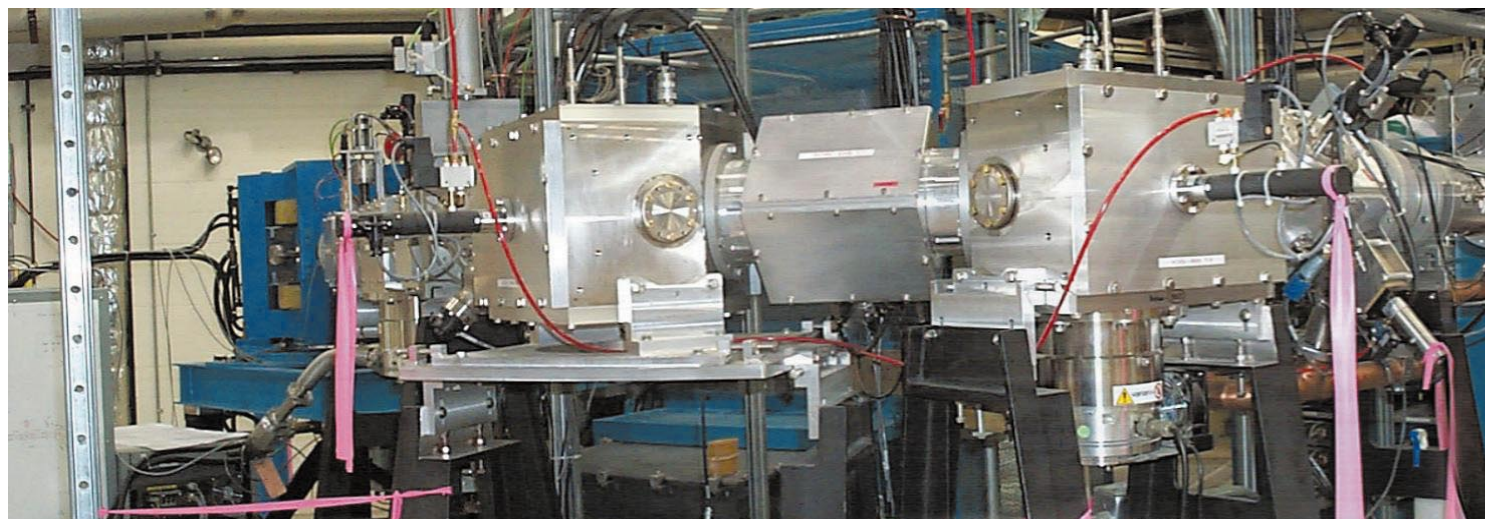
Biosphere 2: future in Earth systems research.

The two-year agreement will seek to create a 'national user facility' at the Biosphere 2 Center outside Tucson. Built in 1991 as a self-sustaining environmental experiment, the centre was turned into a research and teaching campus for Columbia in 1996.

Germany may relax views on biomedical research

Munich Researchers hope that Germany will relax its approach to new biomedical techniques after the appointment of Social Democrat Ulla Schmidt as the health minister.

Schmidt's predecessor, Andrea Fischer of the Green Party, resigned earlier this month following accusations that she mismanaged the BSE crisis in Germany (see *Nature* 409, 275; 2001). Fischer was known for her lack of enthusiasm towards biomedical techniques such as therapeutic cloning.



Bringing supernovae down to Earth

Nuclear physicists, accelerator physicists and astrophysicists are planning a journey into uncharted territory — studying the nuclear processes that occur when massive stars explode. Alexander Hellemans reports.

When the great Russian chemist Dmitri Mendeleev pieced together his periodic table of the elements, little did he know that many of its members were created in huge cosmic explosions called supernovae. Elements lighter than iron can be formed by nuclear fusion in stars. But heavier elements are mostly forged in the extreme environments of some of the most violent explosions in the Universe.

The exact nature of those nuclear processes remains mysterious. Theorists have derived models to explain how unstable heavy nuclei are formed in violent stellar explosions, and then decay into the range of heavy elements that are found in the Universe today. But given the complexity of the processes involved, these models are far from perfect. “Almost every aspect of nuclear physics and particle physics comes to bear in a supernova explosion,” observes Dieter Frekers, a nuclear physicist at Münster University in Germany.

But a series of accelerator projects that aims to bring the nuclear physics of supernovae down to Earth could, within a decade or so, provide the means to test and refine the theoretical models. The accelerators will collide unstable radioactive ions with other nuclei to form a range of exotic nuclei. In charting the wild frontiers of nuclear physics, they could also help astrophysicists better understand the dynamics of supernovae.

As large stars burn up their nuclear fuel and form elements up to iron, they end up with an onion-like structure. Towards the

end of their life, they have an iron core that is surrounded by concentric layers containing increasingly lighter elements towards the outside — silicon, oxygen, carbon, and so on, with helium and hydrogen at the surface. Eventually, the core reaches a mass of about 1.4 times that of our Sun. It then collapses under its own gravity in less than a second.

When the core can collapse no more, it rebounds, and collides with the still-collapsing layers that lay immediately above it. This creates a cataclysmic shock wave that — together with an energetic outpouring of the neutrinos that are formed as protons and electrons in the compressed iron atoms get converted to neutrons — blows away the star’s outermost gas layers in a violent explosion.

In with a bang

During these extreme events, known as ‘type II’ supernovae, nuclei fuse and can also capture neutrons to form highly unstable exotic nuclei that exist momentarily before decaying into more stable isotopes. “The explosion is energetic enough that the shock wave that moves out can, by thermonuclear processes, produce other elements,” says Adam Burrows, an astrophysicist at the University of Arizona in Tucson.

Usually, nuclei made unstable by the addition of a neutron undergo beta decay — the neutron expels an electron and is converted into a proton. The nucleus can then accept another neutron and the process continues. This ‘slow’ neutron capture occurs in certain types of star, and probably also in more violent events such as novae — bursts

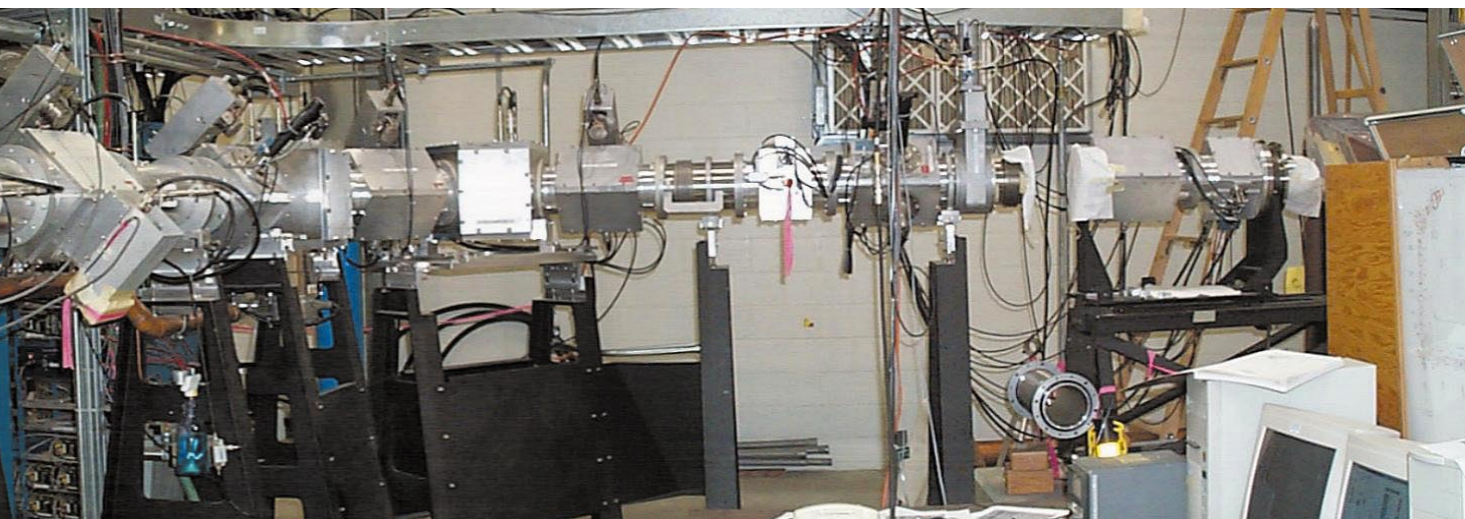
of thermonuclear activity that occur as a white-dwarf star siphons off material from a companion star by gravitational attraction.

But in the intense conditions of a type II supernova, additional neutrons can be added to the nucleus before beta decay occurs. This is called ‘rapid’ neutron capture, and causes highly exotic, unstable nuclei to form. These then decay into familiar heavy elements such as thorium, gold and silver. But the lifetimes of the exotic nuclei produced by rapid neutron capture, and their subsequent radioactive decays, are very poorly understood.

Exotic star performers

If physicists can create the exotic nuclei formed by rapid neutron capture in type II supernovae, they might learn more than simply how the heavy elements in the Universe were made. As much of the violence of a supernova explosion is thought to depend on interactions between the outpouring of neutrinos and the nuclei they encounter, knowing more about the exotic nuclei present should allow theorists to make better simulations of stellar explosions.

Investigating such questions will depend on physicists’ ability to create and study unstable nuclei. Put too many protons, or neutrons, into a nucleus and the excess will immediately be expelled. In theory, nuclei can exist in more than 6,000 different combinations of proton and neutron numbers (see figure, page 450). Less than 300 of these are stable enough to exist naturally on Earth. But using accelerator technology to fragment



Fast track: the ISAC heavy-ion accelerator in Vancouver will experiment with radioactive ions.

stable nuclei, physicists have so far created about 3,000 different unstable nuclei.

These exotic nuclei are produced using two main processes. The first, known as the Isotope Separator On-Line (ISOL) method, was pioneered in the 1960s at a facility called ISOLDE at CERN, the European Laboratory for Particle Physics near Geneva. It involves firing a beam of high-energy protons at a target containing heavy, stable nuclei such as uranium-238, so that some of their neutrons and protons are dislodged. The target is heated to around 2,000 °C, allowing the desired isotopes to escape and be separated using chemical or physical methods. For instance, lasers tuned to specific frequencies can be used to ionize particular elements, which can

then be manipulated using electric fields.

The second method, called projectile fragmentation, is used by several facilities around the world. Heavy nuclei are accelerated up to 60% of the speed of light and hit a thin target of a light material such as beryllium. As they pass through the target, protons or neutrons are stripped from the nuclei, forming a variety of unusual isotopes. Magnetic and electric fields can then be used to separate the isotopes according to their mass, charge and velocity.

Most of the proton-rich nuclei that can exist in theory have already been made, because they are what ISOL and projectile fragmentation tend to produce. As a result, nuclear astrophysicists have been able to

study many of the nuclei created by another process that occurs in extreme astrophysical environments, called rapid proton capture.

Drip fed

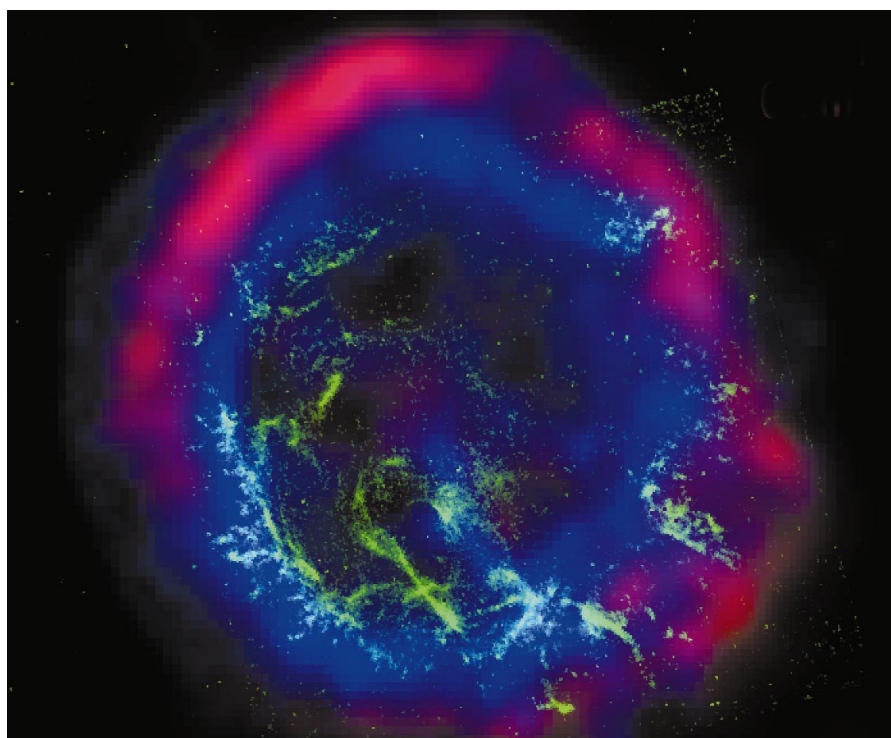
But about 3,000 nuclei that exist between the stable region and the upper limit for neutron-richness — known as the neutron-drip line — remain unstudied. Indeed, the exact position of the neutron-drip line on the landscape of possible nuclei remains unclear for heavier elements. And it is the lifetimes and decays of exotic nuclei in the uncharted territory near this drip line that will provide the key to understanding the formation of heavy elements in type II supernovae.

The challenge is to produce enough exotic neutron-rich nuclei to be able to measure their properties. Not only are they difficult to produce, but they often disintegrate before they can be captured and studied. ISOLDE, for instance, has over the years produced more than 600 isotopes of 68 elements, and can churn out hundreds of billions of exotic nuclei every second. But neutron-rich isotopes such as nickel-78 are produced at rates of only one nucleus every few seconds.

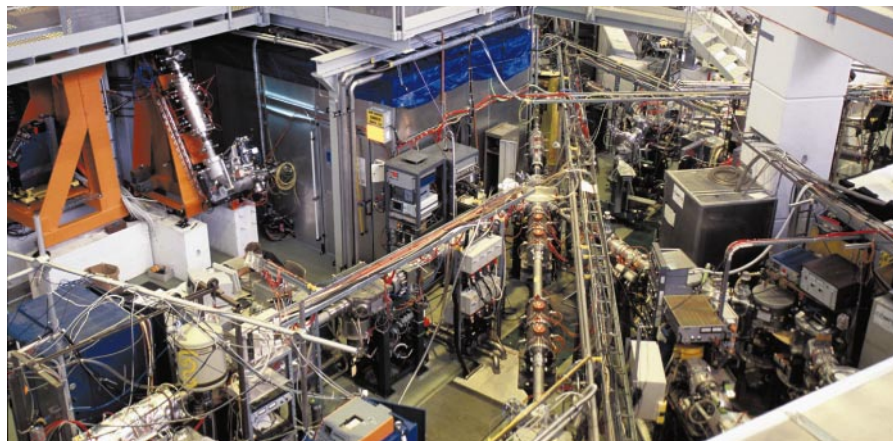
By measuring the lifetimes of these isotopes — which can be shorter than a millisecond — physicists can compare the behaviour of unstable nuclei with theoretical models used to describe nuclear structure. For stable nuclei, these models are well established. The neutrons and protons — collectively known as nucleons — that make up a nucleus are arranged into 'shells'. Within each shell, neutrons or protons have similar energy levels and, according to the laws of quantum mechanics, only a certain number of neutrons or protons can exist within each shell.

A nucleus that consists entirely of full neutron or proton shells is said to contain a 'magic' number of protons or neutrons. Many of these nuclei lie outside the zone of stable nuclei, but they are markedly more stable than their immediate neighbours in the landscape of possible nuclei. Some nuclei are doubly magic — they are magic for both

NASA



Aftershock: the remnant of a type II supernova in which exotic nuclei would have been formed.



Shooting gallery: CERN's ISOLDE fires high-energy protons at heavy nuclei to make unusual isotopes.

proton and neutron numbers — and are more stable still.

When they are formed, magic nuclei persist for longer — up to a few seconds — than other nuclei outside the stable zone. So as neutron capture proceeds in a type II supernova, these magic nuclei will accumulate. “These are milestones in the pathways, and if we are able to find the milestones, we will understand how they are produced — how nuclear synthesis proceeds, step by step, from the iron to the superheavy elements,” says Sidney Gales, director of the Institute for Nuclear Physics in Orsay, near Paris.

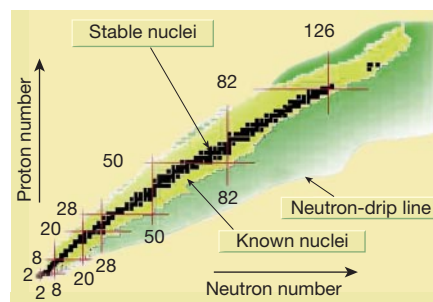
But the problem is that, as you move towards the neutron-drip line, the nuclear shell model seems to break down. “By moving out of stability you are modifying the shell structure in such a way that well-known models do not work any more,” says Gales. Because of quantum-mechanical effects, the loosely bound neutrons in exotic, neutron-rich nuclei do not form clear shells, and the outer neutrons form a ‘halo’ around the core. “This has enormous consequences for the energy level of the nucleons, resulting in different magic numbers,” says Piet Van Duppen of the Catholic University of Leuven in Belgium. It might also help explain the abundances of heavy elements in the Solar System.

Getting heavy

To study neutron-rich isotopes under conditions that are relevant to type II supernovae, it is necessary to accelerate them to high energies. Accelerating beams of unstable, radioactive nuclei was pioneered in 1989 at the Cyclotron Research Centre at Louvain-la-Neuve in Belgium, run by the Catholic University of Louvain, and has been developed further at the Holifield Radioactive Ion Beam Facility at the Oak Ridge National Laboratory in Tennessee, which opened in 1996. So far, the experiments at these facilities have not delved deeply into the exotic territory of heavy, neutron-rich nuclei. But the Oak Ridge facility is now beginning to work with beams of neutron-rich nuclei, such as tin-132 and nickel-56.

Inspired by the examples set at Louvain-la-Neuve and Oak Ridge, other groups are also gearing up to begin accelerating radioactive ion beams. At TRIUMF, Canada's National Laboratory for Particle and Nuclear Physics in Vancouver, a new heavy-ion accelerator called ISAC will in the spring start experiments with a beam of sodium-21 ions, created by the ISOL method. From 2003, an upgrade will allow ISAC to study nuclei with atomic masses of more than 150. At CERN, the ISOLDE facility is now being fitted with a post-accelerator called REX-ISOLDE, which should begin accelerating radioactive beams later this year. Meanwhile, France's GANIL ion-accelerator laboratory in Caen, Normandy, has built a radioactive ion beam facility called SPIRAL. It is essentially a souped-up version of the Louvain-la-Neuve facility, and is currently awaiting permission to start accelerating radioactive ions.

None of the present facilities has the capacity to produce and accelerate the full range of heavy, neutron-rich nuclei produced by rapid neutron capture in type II supernovae. But larger projects now on the drawing-board should provide the means to explore this territory. These ‘second generation’ facilities will produce radioactive beams some 100 times as intense as those that can be achieved today. Their beams will also be more pure, containing precisely defined nuclei with few contaminants. To investigate the rapid neutron-



Combinations of protons and neutrons form a variety of stable (black) and unstable nuclei. The magic numbers of these particles (shown) confer additional stability to the nuclei.

capture process, these facilities will fire neutron-rich nuclei at targets containing deuterium nuclei, which consist of a proton and a neutron. In the resulting collisions, some of the nuclei should gain a neutron.

Projectile splintering

EURISOL, the European ISOL Facility, is currently the subject of a feasibility study being conducted by ten laboratories across Europe. It could be completed — at a site to be determined — by 2010 and will produce exotic nuclei by ISOL. A comparable facility being planned at the GSI National Research Centre for Heavy-Ion Research in Darmstadt, Germany, will accelerate beams produced by projectile fragmentation. Both production techniques will be needed to investigate the full range of neutron-rich nuclei: ISOL has a higher yield, but is limited to isotopes with relatively long half-lives; projectile fragmentation can produce very short-lived isotopes, but has a lower yield. “We are now investigating which targets we will use, which ion sources, which purification methods, and how we will accelerate the ions,” says Van Duppen.

In the United States, the Argonne National Laboratory near Chicago and Michigan State University in East Lansing are bidding to host the Rare Isotope Accelerator (RIA), which could be in operation by the end of the decade. This proposed \$500-million facility will accelerate radioactive ions produced by both projectile fragmentation and ISOL. The RIA should be able to create the majority of theoretically possible isotopes for any element. “What is really exciting and interesting is the ability to reach very far into the neutron-rich terrain,” says Konrad Gelbke of the National Superconducting Cyclotron Laboratory at Michigan State University.

If these second-generation facilities get the go-ahead, nuclear astrophysicists will begin an unprecedented voyage of discovery. Within a decade, predict researchers in the field, exotic beasts that have so far only existed in the thermonuclear cauldron of a type II supernova will finally start to be brought down to Earth. ■

Alexander Hellemans is a science writer in Naples.

ISOLDE

♦ <http://isolde.web.cern.ch/ISOLDE/welcome.html>

Louvain-la-Neuve cyclotron

♦ <http://www.cyc.ucl.ac.be>

Holifield Radioactive Ion Beam Facility

♦ <http://www.phy.ornl.gov/hribf/hribf.html>

TRIUMF-ISAC

♦ <http://www.triumf.ca>

GANIL-SPIRAL

♦ <http://ganinfo.in2p3.fr/spiral>

EURISOL

♦ <http://www.ganil.fr/eurisol/index.html>

GSI Future Facility

♦ <http://www.gsi.de/GSI-Future>

Rare Isotope Accelerator

♦ <http://www.phy.anl.gov/ria>

The brain in Spain

The legacy of Santiago Ramón y Cajal, the brilliant Spanish neuroscientist, is to be preserved in a new museum. But the fight to recover his lost works goes on, say Xavier Bosch and Alison Abbott.

For a boy who dreamed of being an artist, but started his career apprenticed to first a barber and then a cobbler, Santiago Ramón y Cajal made a distinguished mark in science. He was both Spain's first Nobel laureate and a founding father of modern neuroscience. But since Cajal's death in 1934, his work — including a priceless archive of drawings — has been consigned to basements, maltreated and given away as gifts.

Now Cajal's legacy is to get a permanent home in Madrid. Spain's national research agency, the CSIC, and the philanthropic Marcelino Botín Foundation agreed last month to co-sponsor the building of a Ptas130 million (US\$740,000) museum to house Cajal's collected works. It should open in 2002, the 150th anniversary of Cajal's birth.

Cajal shared the Nobel Prize in Physiology or Medicine in 1906 for his theory that the nervous system was made up of individual cells, later termed neurons. Until Cajal, most researchers supported the 'reticular' theory, believing that the system was a network in which cells fused to give long, continuous nerve fibres. As talented a draughtsman as he was a scientist, Cajal depicted the synapses that separate individual neurons in meticulous drawings of histological preparations. These rival today's images from electron and confocal microscopy.

Even more remarkable was Cajal's uncanny ability to infer function from observation. "We are still working on the theories that Cajal developed," says Marina Bentivoglio of the University of Verona. "He had tremendous insight and could understand what he saw much better than anyone, even today." Barry Everitt of the University of Cambridge, editor of the *European Journal of Neuroscience*, agrees: "He has rarely been shown to be wrong in his predictions about neuronal connections."

Cajal's will — handwritten one month before his death — stated that his works should be preserved at his institute in Madrid. Shortly afterwards, the collection was taken from his home to the Cajal Institute. But they were not well curated. At the end of the Spanish Civil War in 1939, they were found among rubble in the institute's basement. And although a small museum displayed a sample

of Cajal's works from 1945 to 1985, the majority remained stored in boxes.

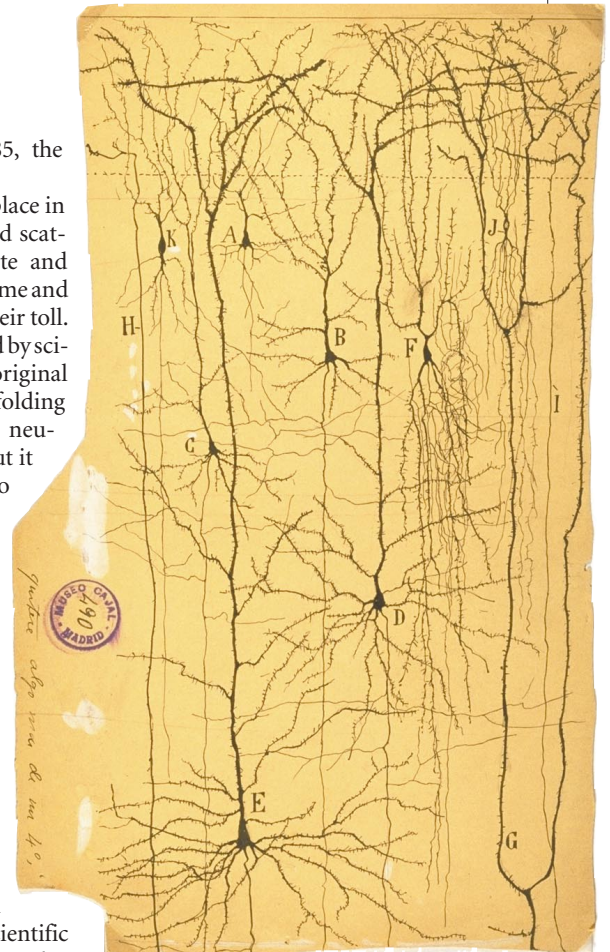
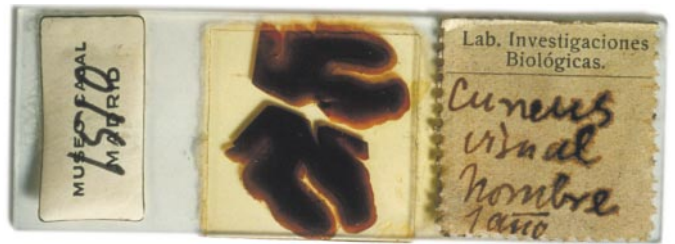
The first general inventory took place in 1975, when Cajal's works were found scattered around offices at the institute and Madrid's Complutense University. Time and poor storage conditions had taken their toll. "The most serious damage was caused by scientists sticking adhesive tape onto original documents, writing on them or folding them," says Miguel Freire, a neuroanatomist at the Cajal Institute. But it was not until 1997, when Ricardo Martínez Murillo became director of the institute, that a concerted effort to preserve Cajal's legacy began.

Martínez Murillo appointed Freire to head the Cajal Legacy Team. The team is now making a detailed computerized inventory of Cajal's works and organizing restoration. The collection is thought to comprise about 2,000 scientific drawings, 2,300 histological preparations and 2,100 scientific and personal letters, as well as portraits, photographs, notebooks, laboratory materials and personal objects. If the collection can be turned into an accessible scientific resource, says Everitt, it would be very valuable. But about a quarter of Cajal's works are thought to have been lost.

In the past, says Cajal's grandson, Santiago Ramón y Cajal Junquera, a neuropathologist at Zaragoza University, it was common for scientific visitors to the Cajal Institute to be given a Cajal drawing as a gift. Now that the museum is in prospect, Freire and Martínez Murillo hope that those who hold Cajal artefacts will return them to the collection — they plan to issue an appeal over the Internet once the inventory is complete.

Carlos Belmonte, director of the neurosciences institute at Miguel Hernández University in Alicante, and his colleague Roberto Gallego both have Cajal drawings, given to them by Gallego's father, a pupil of Cajal's. Both say that they are prepared to discuss loaning the drawings to the new museum.

Members of Cajal's family back the appeal to return artefacts to the museum, although they are in dispute with the Cajal Institute over who actually owns them. These ques-



Networking: Cajal translated histological preparations (top) into intricate drawings.

tions might have been resolved two years ago, when Fernando de Castro, an heir of one of Cajal's pupils, attempted to auction his collection, and the family claimed ownership of the works. A judge blocked the sale, but did not rule on whether the artefacts belonged to the family or the institute.

The poor cataloguing and curation of the collection in Madrid may have deterred some people from returning their Cajal artefacts. The museum's supporters hope that the latest developments change their minds. "I urge all to participate in this process," says Charles Ribak, a neuroanatomist at the University of California, Irvine, and president of the Cajal Club — the oldest neuroscience organization in the United States.

Xavier Bosch writes for *Nature* from Barcelona; Alison Abbott is *Nature's* senior European correspondent.

♦ <http://www.cajal.csic.es/legadoca/legadoin.htm>

European researchers encourage improvement in lab animal welfare

Sir—Rodents are currently excluded from the protective regulations of the US Animal Welfare Act. Last September the US Department of Agriculture (USDA) decided to take steps to include rats, mice and birds in the act. This decision was heavily criticized by the Association of American Medical Colleges, the National Association of Biomedical Research and the Federation of American Societies for Experimental Biology^{1,2}. Yet the USDA is determined to implement its decision and is planning to make it effective by October.

The opposition by these leading science organizations in the United States is not supported by the American Association for Laboratory Animal Science and also contrasts with European perspectives on the issue. Last September, the European Science Foundation (ESF) published a position paper on the use of animals in research³. The guidelines from this paper, summarized below, were adopted by the ESF assembly at the end of November.

The ESF consists of 67 leading science-funding agencies, research councils and academies of science from 23 European countries. Its role is to stimulate cooperation between national organizations and individual scientists from different countries and to advise on science policy. Its position paper states the conditions that must be met to make the use of animals for research purposes morally acceptable. The ESF encourages scientific organizations and individuals involved in animal experimentation to follow these guidelines, which can be summarized as follows:

- Laboratory animals not only have an instrumental value, but also an intrinsic value, which must be respected.
- While accepting the need for animals to be used to advance scientific knowledge and to promote human and animal health and well-being, the ESF strongly endorses the reduction, replacement and refinement principles.
- Research to improve the welfare of animals should be encouraged and actively supported.
- Before a programme of research, the proposed animal use should be evaluated independently, including an assessment of the likely benefit and suffering.
- Investigators should assume that procedures that would cause pain in humans also cause pain in other vertebrates, unless there is evidence to the contrary.
- Investigators should be adequately educated and trained through accredited courses on laboratory-animal science,

including discussion of animal alternatives, welfare and ethics.

- The ESF encourages journals to include in their publication policy a statement on the ethical use of animals.

The ESF recognizes the important role of journals in the ethical use of animals in research. It is, therefore, unfortunate that most journals publishing in this area do not have stated guidelines that must be a prerequisite for consideration of manuscripts⁴. A recent study shows that only 13 of 83 relevant journals contain a comprehensive statement on the ethical use of animals. In 42 journals (59%) there was no statement at all⁵.

The ESF guidelines demonstrate that the scientific community in Europe is leaving behind its defensive attitude, and is more proactive in expressing what it sees as standards to adhere to when it comes to the use of animals in research. This should be a benefit for animals and for science.

Bert van Zutphen

Department of Laboratory Animal Science, Faculty of Veterinary Medicine, Utrecht University, PO Box 80.166, 3508TD Utrecht, The Netherlands

1. *Nature* **407**, 659 (2000).
2. Hendrix, M. *Nature* **408**, 133 (2000).
3. <http://www.esf.org/ftp/pdf/SciencePolicy/ESPB9.pdf>.
4. Hampshire, V. *Nature* **407**, 671 (2000).
5. van Zutphen, L. F. M. & Festing, M. F. W. in *Progress in the Reduction, Refinement and Replacement of Animal Experimentation* (eds Balls, M., van Zeller, A.-M. & Halder, M. E.) 1375–1382 (Elsevier, Amsterdam, 2000).

Users must help to keep public databases correct

Sir—With the continued growth of the public DNA sequence databases, and the recent addition of the 11,000,000,000th nucleotide to GenBank (including DDBJ, EMBL and GenBank), it is timely to assess how we use these databases.

GenBank is the archive of all publicly available DNA, RNA and protein sequences. Upon publication a new sequence and its annotations appear in it. Investigators use GenBank in many ways, most commonly for similarity searches such as BLAST; to retrieve records; and for sequence analysis, multiple sequence alignment or pattern finding search. Errors sometimes occur in GenBank, ranging from the trivial (incorrect postal codes), to the misleading (30 nucleotides of vector left on the ends of a record), to the mission-critical (a full length mRNA without a coding sequence (CDS) annotated on it). Also very common are incomplete references that prevent researchers from linking the GenBank record to the publication that refers to it first.

Over the years some people have chosen to report these errors, but in most cases

they are left unmodified. An uncorrected 'discovered' error is one of the worst possible failings in GenBank, so if you discover an error, report it to the database (update@ncbi.nlm.nih.gov) and it should be rectified — although a follow-up is advised to make sure this gets done.

If you are a submitter, look at the record you submitted a few years ago: is it still correct? Was the citation ever updated? Take pride in the sequences that carry your name! Our ability to interpret genomes depends on all of these records being as accurate as possible. This is a task for all users of the databases.

Francis Ouellette

Bioinformatics Core Facility, Centre for Molecular Medicine and Therapeutics, University of British Columbia, 950 West 28th Avenue, Vancouver, British Columbia V5Z 4H4, Canada

The long-term answer: fight fire with research

Sir—We agree with your Editorial (*Nature* **406**, 661; 2000) in that the spectacular fires of last year should be a stimulus for improving management of forest ecosystems in the American west. In fact, through the leadership of the US Department of Agriculture Forest Service, a new programme is being implemented to improve forest management, in view of fire and other hazards that endanger forest health.

However, we strongly disagree with your statement that the Forest Service "uses its money to pay its staff, not to conduct high-quality research into forest management".

In reality the Forest Service, beginning in the 1920s under Rafael Zon, has exerted global leadership in forestry research. The ecosystem models used to manage fires in the west and to accomplish appropriate restoration are products of long-term ecological research, started decades before the value of such research was realized in the academic community or other scientific organizations. Five of the National Science Foundation long-term ecological research sites are Forest Service research sites, where such work has been taking place for many years.

Forest Service research is cost-effective. About 7% of the agency's budget, some \$220 million a year, supports six research stations, the Forest Products Laboratory and the International Institute of Tropical Forestry.

Thomas J. Mills*, Ariel E. Lugo†

**Pacific Northwest Research Station, PO Box 3890, Portland, Oregon 97208, USA*

†International Institute of Tropical Forestry, PO Box 25000, Rio Piedras, Puerto Rico 00928-5000

Memoirs of a reluctant cult figure

A sceptic about conventional wisdom and a seeker of unity in nature.

Homage to Gaia: The Life of an Independent Scientist

by James Lovelock

Oxford University Press: 2000. 396 pp. £19.99

Crispin Tickell

James Lovelock has been called many things in his long life, from “mad inventor” to “holy fool” (although, as he remarks, “wholly fool” may have been intended). No wonder he has become something of a cult figure. This is a matter of embarrassment to him, for the effect is to caricature or at least give a wrong impression of ideas that, under his and other auspices, are now entering mainstream science. A prime example is the Gaia hypothesis, a metaphor no more or less allusive than others such as “the selfish gene”.

Lovelock's place as an inventor and scientist emerges clearly from this autobiography. His curiosity about the natural world was dominant from the start. Fortunately, a schoolboy experiment tempting girls from another class to eat deadly nightshade berries (as he calls it, an apprenticeship to the Borgias) was stopped in the nick of time. His father, no mean inventor and household-fixer himself, gave him a wooden box filled with electrical odds and ends for Christmas when he was four, and this opened opportunities not just for experiment but for far-reaching questioning about how and why things work. Lovelock remembers his early years as full of happiness and sunshine.

Now, as in the 1960s, science is practised by the big battalions, and research teams are composed of specialists. Lovelock has always had ideas far beyond any speciality, although he has contributed to many such. It was the breadth of his range that prompted the sensational arrival, apparently out of the blue, of a letter from NASA in March 1961, inviting him to join a group of scientists whose mission was to explore the surface of the Moon. With some misgivings he left the Mill Hill Institute in London, and began, first in the United States and then in England, what he describes as the independent practice of science, which he has continued to this day.

In Britain as elsewhere, there are, as Lovelock says, internal tribal barriers between the sciences (particularly marked, for example, between physics and biology). Overcoming them was perhaps easier in the United States, and Lovelock found undreamed-of opportunities, which he exploited on his return to Britain years later. He was thus able to act as adviser to such companies as Shell, to Lord Rothschild, and to many others, and to exercise his remarkable inventiveness.

The list of his inventions, most based on

laboratory work with relatively simple equipment, is impressive. Beginning with the argon detector (designed to analyse compounds from a gas chromatograph column), he went on to create a device for measuring the intensity of heat radiation, a transmodulator, probably the first working microwave oven and, most famously, the electron capture detector. This last invention led to the discovery of ozone depletion by chlorofluorocarbons and the series of international agreements to end their manufacture and use. The stories behind these inventions form one of the most fascinating parts of the book. It was the same kind of reasoning that enabled Lovelock to tell sceptical NASA engineers that it was unnecessary to send the Viking mission to Mars to show that there was no life there today. The composition of the martian atmosphere, detectable from Earth, showed that it was a dead, high-entropy planet close to chemical equilibrium.

This insight led to another of greater significance. The Gaia hypothesis was born of the realization in September 1965 that, although the Earth's atmosphere was in chemical disequilibrium, it was nonetheless stable and had been so for billions of years. Something was therefore regulating the atmosphere. If most atmospheric gases came from living organisms, then it must be life at the surface that did the regulating through a complex system of feedbacks. In other words, the Earth was behaving like a living organism. It was the novelist William Golding who suggested Gaia to Lovelock as the name for a self-regulating Earth.

The Gaia hypothesis developed over the years. In 1974 it was described by Lovelock and the microbiologist Lynn Margulis as “a tightly coupled system whose constituents are the biota and their material environment, which comprises the atmosphere, the oceans and the surface rocks”. To illustrate the means by which feedback mechanisms maintained stability, Lovelock invented a theoretical model which he called Daisyworld, in which black daisies and white daisies acted to their own selective advantage as a kind of planetary thermostat. The model has been much elaborated in response to criticism, but its underlying principles remain the same.

Lovelock gives a good description of the initial reactions to the Gaia hypothesis. Chemists, physicists and engineers, familiar with nonlinear systems, found no difficulty in accepting it. But many biologists did not like it at all, although some, notably the late W. D. Hamilton, later became converts. For those who took any notice, it smacked too



Lovelock at work on his electron capture detector, which identified atmospheric ozone depletion.

much of the age-old suggestion that the Earth was itself alive, and Gaia was clearly not a living organism in the usual sense of the term. Lovelock himself was a little careless in his use of language. As he says, his first book on the subject was “more a love letter than a text book”. Moreover, it proved alarmingly attractive to those in search of a new religion, who saw the Gaia hypothesis as a kind of legitimization of a mother-goddess. It has taken time, and careful research (some of it published in *Nature*), to establish Gaia's scientific credentials as Earth systems science. This fits admirably with the objectives of the Geological Society of London, with which the Gaia Society is at present being merged.

Lovelock's autobiography brings out his single most important characteristic as a scientist: his refusal to accept dogmas in small things or in big, or to establish dogmas of his own. His skill as an inventor and his insights as a scientist come from scepticism about the conventional wisdom and from his ability to bring together somewhat unlikely elements from different disciplines — hence his suspicion of big science, his feelings about the sometimes stifling character of peer review, and his belief in the individual approach.

As for dogmas of his own, Lovelock has

ANTHONY HOWARTH/SEPL

book reviews

always maintained that his ideas are provisional. As I know from personal knowledge of him, he hates the polarizations that public controversy brings. A profound agnosticism runs through his book. With it goes a painful honesty leavened by *joie de vivre*, manifest even in dialogue with his computer. Indeed, he finds composing computer programs like writing poetry, with the blame for any mistakes invariably his own.

Almost anyone with an interest in science will enjoy this book. If I have a criticism, it is that the chronology is occasionally confusing, so that the ways in which the author and his thinking evolved get lost. The story is one of a very human being whose mind grows in the play of experiment and ideas. Like the philosopher Mary Midgley, he believes that one of the blind spots in current society is its alienation from the world around it. For him, Gaia represents the essential unity of the living and physical environment. We do not have to be holy — or wholly — fools to recognize it. ■

Crispin Tickell is at Ablington Old Barn, Ablington, Cirencester GL7 5NU, UK.

Floral tributes for a treasure reclaimed

A Garden for Eternity: The Codex Liechtenstein

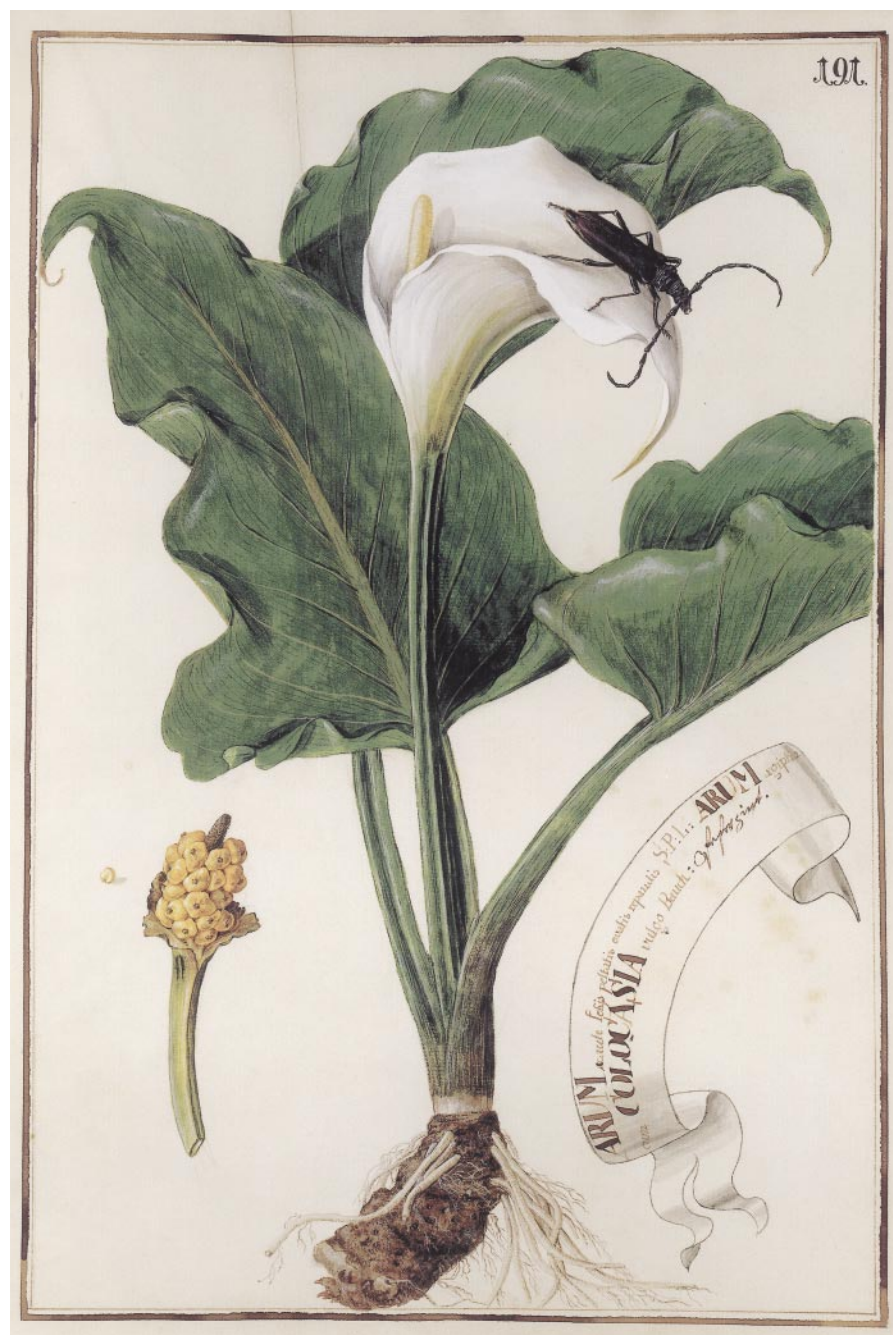
by H. Walter Lack

Bentley: 2000. 344 pp. \$140

Ryan J. Huxtable

Botanical illustration has a long history. The painted garden, or *hortus pictus*, was supremely important for the thousands of years when nearly all medicinal drugs were derived from plants, and there was no effective technology for preserving herbarium specimens from decay, attack by insects or other destructive agents. Accurate depictions were, therefore, essential for teaching, identification and preparation of medicines.

The Royal Botanic Gardens at Kew in London have attracted many fine resident artists. Among the greatest were the Bauer brothers, Franz and Ferdinand. Although the brothers spent most of their working lives at Kew, Franz working there for more than 50 years, they came from Feldsberg, then in the Hapsburg Austrian Empire. The town is now called Valtice, and is in the Czech Republic, on the border with Austria. Here, between 1770 and the mid-1780s, at the beginning of their careers, the Bauer brothers illustrated the first volumes of the 14-volume manuscript now known as the *Codex Liechtenstein*. This work was the creation of Norbert Boccia (1731–1806), prior of the Monastery of the Brothers of Mercy in



Blooming beauty: a water colour of *Zantedeschia aethiopica*, produced by the Bauer brothers.

Feldsberg. The monastery ran a hospital, and Boccia developed a medicinal garden that he used for teaching and for growing plants to obtain drugs. The *Codex* was intended as a work of instruction and reference for the monastery. It took 34 years to complete, by which time the Bauer brothers had long gone and other, lesser hands had taken over the illustrations. The final *Codex* contained 2,748 watercolour paintings of plants, each 35 by 30 centimetres, and each taking about a week to complete.

A Garden for Eternity reproduces 88 of these plates, in large format. Sixty-three of the plates are by the Bauer brothers, working either alone or together. The paintings by the brothers are of such high artistic and

botanical quality that they belie the editorial comment that “for a scientific illustrator there can be no artistic freedom, he must work like a camera”. Accuracy and artistry are not incompatible. Typically, the plants are aesthetically arranged without compromising their natural appearance. The foxglove, *Digitalis purpurea*, was painted in 1775, ten years before the appearance of William Withering’s book, *An Account of the Foxglove*. The frontispiece of Withering’s book reproduces a painting of the foxglove taken from *Flora Londinensis*, by the English eighteenth-century botanist William Curtis. The Bauer and Curtis figures are of equal botanical accuracy, but the Bauer foxglove is disposed in a way that elevates the work from illustration to art.

As a further artistic touch, the name of each plant is shown in an individualized ornate, scroll-like frame. The Bauers celebrate the glory of the commonplace — peas, grapes, carrots, beets — with the occasional insect, and in one case a house sparrow, enlivening their plates.

If a necessary attribute of a coffee-table book is that the illustrations are more important than the text, it would be unfair to use this epithet for *A Garden for Eternity*. But the text is idiosyncratic. It is partly a general history of botanical discovery in the eighteenth century and partly a collection of biographies of people associated with the *Codex*.

There is more detail about various other figures, though, and a touch of the theatrical approach with the inclusion of a “dramatis personae” and the description of Boccus as the book’s “second hero”. Its first hero is Nikolaus Joseph Jacquin. Jacquin, from Leiden in the Netherlands, was a noteworthy figure in medicine, chemistry and botany, and an exemplar of the holism of education and culture that has been largely lost in our age of specialization. In addition to many works on botany, Jacquin wrote a treatise on pharmacology. He was also a trained artist, producing more than 2,000 coloured copperplates during his career.

Although an important figure in Austria, being a professor of chemistry and botany at the University of Vienna and director of the university’s botanical gardens, Jacquin does not seem to have been particularly closely

involved in the *Codex*. He is described in *A Garden for Eternity* as “probably” being an adviser in its creation. But his biography is interesting. His botanical explorations took him to the Caribbean, where he named the mahogany tree. Plants collected over a four-year period on this trip (1755–59) were destroyed by ants, and so Jacquin turned instead to painting *in situ*, laying the foundations of his book, *The Natural History of Selected American Plants*.

As the *Codex* neared completion, the ageing Boccus presented it to the Prince of Liechtenstein in 1799. In exchange, the prince gave his support to the monastery’s efforts to help the sick and poor, as the Liechtensteins were major landowners in the Feldsberg area. The *Codex* was kept in the Liechtenstein library in Vienna for 150 years. However, with the approach of Soviet troops towards the end of the Second World War, it was first moved to a mine in the Salzkammergut and then smuggled into Vaduz in Liechtenstein, where it remains.

A Garden for Eternity is the first partial publication of the *Codex*’s contents. It is a contribution to botany, art and history, and introduces a long-hidden treasure to the world. Further information on the *Codex Liechtenstein* (in German) is available at <http://www.bgbm.fu-berlin.de/bgbm/research/data/lack/>.

Ryan J. Huxtable is in the Department of Pharmacology, College of Medicine, University of Arizona, Tucson, Arizona 85724-5050, USA.



The science behind the dish: fine cooking can reap gustative rewards from chemistry.

New in paperback

The Promise of Sleep: A Scientific Connection Between Health, Happiness and a Good Night’s Sleep

by William C. Dement with Christopher Vaughan
Pan/Dell, £6.99/\$14.95

Natural Capitalism: Creating the Next Industrial Revolution

by Paul Hawken, Amory Lovins & L. Hunter Lovins

Back Bay Books/Earthscan, \$17.95/£12.99

“A visionary book that is truly appropriate for the new millennium with its focus on ultra-efficient use of energy and materials.” Norman Myers, *Nature* 402, 13–14 (1999)

Time, Love, Memory: A Great Biologist and His Quest for the Origins of Behaviour

by Jonathan Weiner

Faber & Faber, £8.99, \$14

“Weiner’s book is well written and fun to read, although the question arises, who is the audience? This is scientific reportage and as such should not be expected to provide in-depth

analysis to tenets and conclusions. Therefore, those who wish to find a critical assessment of neurogenetics should look elsewhere.”

Yadin Dudai, *Nature* 398, 773–774 (1999)

Principles of Condensed Matter Physics

by P. M. Chaikin & T. C. Lubensky
Cambridge University Press, £29.95, \$47.95

Ants at Work: How an Insect Society is Organized

by Deborah Gordon

W. W. Norton, \$13.95, £9.95

“*Ants at Work* is an entertaining mixture of personal travel reports and popular science.” Jürgen Heinze, *Nature* 401, 856–857 (1999)

Murder, Magic and Medicine

by John Mann

Oxford University Press, £9.99, \$19.95

Tadpoles: The Biology and Anuran Larvae

edited by Roy W. McDiarmid & Ronald Altig
University of Chicago Press, \$40, £25.50

From chemistry to cordon bleu

The Science of Cooking

by Peter Barham

Springer: 2001. 244 pp. £19.95, \$34.95

Etienne Guyon

Nicholas Kurti, a physicist from Oxford who died in 1998, was better known for his low-temperature work on nuclear magnetism than for his interest in fine cooking or his role in promoting molecular gastronomy as a respectable field of scientific investigation. Nevertheless, the science behind cooking fascinated him, and he used to say it was paradoxical that we should know more about the temperature of a star than of a soufflé.

Kurti’s enthusiasm was, writes Peter Barham, one inspiration behind this book. And just as scientists have embraced cooking, so some fine cooks over the years have concerned themselves with the scientific aspects of the culinary art, including the eighteenth-century adventurer Count Rumford, court patissier Marie Antoine Carême in the early nineteenth century, and the celebrated gastronome ‘Curnonsky’ half a century ago. The interest remains alive today and is exemplified by *The Science of Cooking*.

Given the gustative rewards that result from understanding the science of cooking, and its importance to the food industry, this keen interest is not surprising. For readers of *Nature*, developments in the physicochemical study of what is known as 'soft condensed matter' (in France we use the less attractive expression *matière molle*, or sometimes *objets fragiles*) are an added incentive to explore this area. Although the three expressions imply different realities, they actually correspond clearly to various states of food.

The Science of Cooking will also be stimulating for amateur cooks with an interest in following recipes and understanding how they work. They will find anecdotes and, sprinkled through the book, scientific points of information, such as how the colour of fish meat is determined by the density of the living fish relative to that of the water it lives in.

I had, however, a few concerns about the book. The first is generic and unavoidable. The book is a little short on basic explanations of such subjects as interfacial and colloidal physics, polymers and gels, and granular matter, which are subtle physical ingredients of cooking science. The intricate chemistry of cooking — for example, the famous Maillard reactions between amino acids and carbohydrates that give meat its taste — is not explained in any depth.

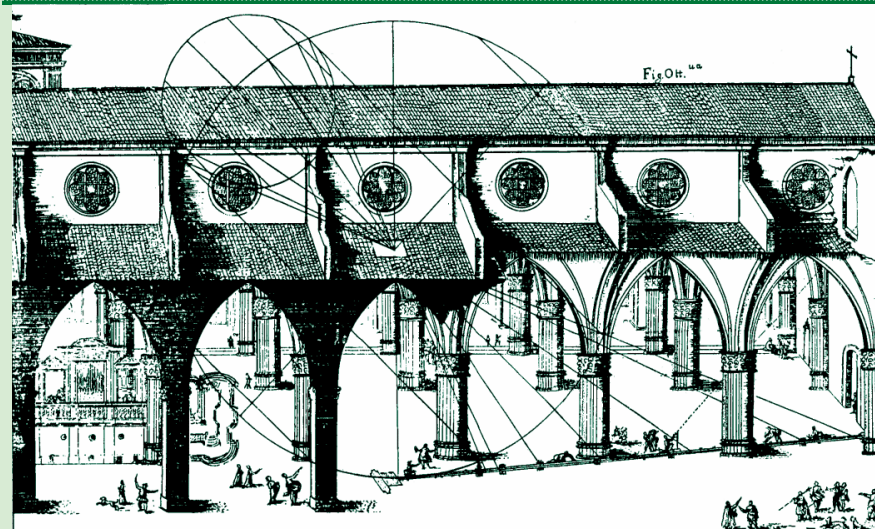
On the other hand, detailed descriptions might have been too technical for the wide audience the author is writing for. One possibility would have been to use illustrations alongside the descriptions given in the book — pictures are, in fact, absent except in the introduction. Personally, I prefer cookery books to include photographs of the recipes.

The choice of recipes is also rather traditional and does not do justice to the present interest in the science of cooking. Restricting a chapter on sauces to a vinaigrette and a mayonnaise is too limiting. I would have at least liked to read about a Béarnaise or a hollandaise, given the different type of emulsion (coated butter) involved in these sauces and the associated subtle stability problems. Cooking meat at a fairly low temperature is an interesting alternative to the high temperatures of grilled meat and deserves more attention; it is the secret of an Argentinian *asado* as well as of a *braisé*. Finally, for a book that presents itself as a science book, the lack of references is a significant omission.

Despite these remarks, the book is a pleasant read and is an invitation to become better acquainted with the science of cooking. ■

Etienne Guyon is at the École Normale Supérieure, 45 rue d'Ulm, 75005 Paris, and at the Université Paris-Sud, 91405 Orsay, France.

Science in culture



Cassini's craft

Contributions of a seventeenth-century astronomer.

Martin Kemp

Spectacular images of Jupiter's 'red spot' — assembled from data sent back by the latest space probe to skirt the planet — made the pages of many newspapers in the first days of 2001. Such probes have regularly been named after famous explorers of the heavens and Earth. Galileo is still orbiting Jupiter after five years, and Magellan let us look afresh at Venus in the early 1990s. By contrast, the name assigned to the new joint NASA–European probe, which is due to rendezvous with Saturn in 2004, rings a less obvious bell in the popular mind. How many people, outside the world of astronomy, could say what Cassini has done to merit such an accolade?

In fact, the choice of name is unusually apposite. Gian Domenico Cassini was renowned for his exceptional mastery of the astronomer's subtle art of educated seeing. He exploited the new generation of giant refracting telescopes to enlarge our knowledge of both Jupiter and Saturn. Cassini provided tables of the motions of Jupiter's satellites, formulated rotation periods for Jupiter and Saturn, established the oblate shape of Jupiter, and observed Saturn's surface belts and four of its satellites. In 1675, six years after his move to the splendid new observatory in Paris, he discovered the "Cassini Division" in the ring of Saturn. He also played a valuable role in the confirmation that the planet's 'ears' corresponded to a ring.

Astronomy at the highest level was then, as now, a matter of international prestige and big money. Scale was crucial if precise observations were to be made — one of the lenses sent to Cassini's Parisian observatory boasted a focal length of 136 feet. One factor, however, that distinguished the patronage of astronomy in Cassini's time from that of our era was the role of the Church. Notwithstanding the unhappy episode of Galileo's forced retraction of his

The Bologna Meridian from G. D. Cassini's *Meridiana* (1695).

Copernicanism, the Church generally found astronomers' revelations of the divine geometry of the heavens much to its liking.

The Church did more than provide an intellectual and theological base for astronomers during the sixteenth and seventeenth centuries. For example, the great cathedrals and basilicas, which enclosed by far the largest interior spaces, could be set up as massive scientific instruments, most notably as locations for extended meridian lines. These numerous church meridians attracted considerable public interest as well as serving their functions of time-keeping, astronomical recording and calendar reform.

Notable among the church meridians is that found in the city of Bologna, where the devout Cassini first came to prominence. The city acquired its first meridian in 1576, when Egnatio Danti opened a small aperture in a chapel wall of the massive late Gothic basilica of San Petronio, and inlaid a line across the floor in order to track, as closely as was practicable, the direction of the midday sun as it slanted across the tiled pavement. In 1655, Danti's meridian was replaced by Cassini's meticulously engineered device, its inset metal strip aligned due north and passing diagonally with breathtaking precision between the massive piers of San Petronio. Bernard de Fontenelle hailed the feature as "a new oracle of Apollo where the sun can consult with confidence about all the difficulties of astronomy".

For his part, Cassini promised the "illustrious nobles" in charge that "the kingdom of astronomy is now yours". Restored on a number of occasions, not least by the astronomer himself in 1695, its enduring beauty provides testimony to his importance, which is now appropriately recognized through the 'new' Cassini. ■

Martin Kemp is in the Department of the History of Art, University of Oxford, 59 George Street, Oxford OX1 2BE, UK.

The suspense of strangeness

Science supplies poetry with a register of words outside common usage.

Maurice Riordan

There is a line about moonlight by Sylvia Plath that I find haunting: "This is the light of the mind, cold and planetary." My ear is arrested by the rhythm — by the way the line runs quickly into the strong stress on "mind", which then collides immediately with the stress on "cold" before the line spins off into space. Except the space created by the metaphor is inner space, so that there is also a collision of inner and outer worlds. The effect is disturbing, is meant to be disturbing: the imagining mind is breached, cleaned out and (for a moment) exhilarated by the hostile image of nature.

The line is scientific, of course, only in the flimsiest sense. Even so, to the degree that it is, it is crucial to the effect: "light" generally brings associations of illumination, vision, transcendence. But not here: the use of "planetary" erases these associations and replaces them with physical, actual light.

This illustrates something about how poets use words with scientific associations. They exist in a state of sensitive relationship with everything else in a poem — and on an equal basis, which means they have to work: they must earn their keep by at least increasing the surface tension of the language. They are, then, part of the poem's imaginative exertion, that promiscuous agility of language which, as described by the nineteenth-century Italian poet Giacomo Leopardi, "keeps the mind in constant and lively movement and action, transporting it suddenly, and often abruptly, from one thought, image, idea, or object to another, and often to one very remote and different; so that the mind must work to overtake them all, and, as it is flung here and there, feels invigorated".

One very remote and different: Leopardi's phrase suggests the particular lure science has for poets. Science supplies a register of words outside common usage; it provides a stock of ideas often far removed from our everyday observation of things; and this body of knowledge seems to be an objective description of the world. Thus, science has the potential to give a poem what the American poet Elizabeth Bishop called "the suspense of strangeness", a sense of being in a world that is at once engrossing and unpredictable.

This rules out attempts to write earnest, educative poems about science. Scientists are

sometimes puzzled by this. Why doesn't a scientific sense of wonder translate readily into poetry? To give an adequate answer might require looking back to classical poetry, when a religious and transcendent view of nature dominated the imagination. But my aim is more reachable: to cite some well-known English poems that seem to me to exemplify how poets can engage imaginatively with words and ideas that derive necessarily from a materialist model of the world.

This is John Donne, whose life coincided with the beginnings of modern science and whose great theme is the body spiritualized and eroticized by love, writing just when William Harvey was pioneering anatomical research:

*Know'st thou but how the stone
doth enter in
The bladder's cave and never break
the skin?
Know'st thou how blood, which to
the heart doth flow,
Doth from one ventricle to th'other go?
And for the putrid stuff which thou
dost spit,
Know'st thou how thy lungs
have attracted it?*

Never mind that Donne opposes scientific enquiry of any kind. This does not limit the daring of the lines. They imagine — for the first time, I think, in English poetry — the body as a biological entity. They take us from the outside to the inside, visualizing the body in familiar physical images but also using technical terms. This elasticity creates a tension between the personal and the impersonal, the body as 'thou' and the body as object. In fact, the whole passage, some 40 lines from 'The Second Anniversary', conveys both a religious despair at human ignorance and a transgressive curiosity about our physical nature.

Science is an unfolding narrative, one that answered Donne's questions only to raise a good many more. It posits the world as discoverable, perhaps indefinitely so. But it also reveals nature as different. It produces a version of reality at odds with orthodox thinking, and one that in any case is often beyond human measure in its scale and complexity. Few great poems in English open their doors to this northern wind, and when they do they are hinged with reluctance. Alfred, Lord Tennyson, like most nineteenth-century poets, was appalled by the



Poet Elizabeth Bishop believed some words could create a world at once engrossing and unpredictable.

Science produces a version of reality at odds with orthodox thinking.

mechanistic implications of contemporary science; and he rejected them in his elegiac sequence 'In Memoriam': "I think we are not wholly brain, Magnetic mockeries ..."

*Not only cunning casts in clay:
Let Science prove we are, and then
What matters Science unto men,
At least to me? I would not stay.*

To be appalled, I'd suggest, is neither unpoetic nor unscientific. Tennyson had absorbed Charles Lyell's *Principles of Geology*, and, elsewhere in the poem, we see how the concept of deep time broadens the perspective of his mourning. It also intensifies the style, so that Leopardi's words come to mind for the pleasurable swift alternation of images, for the seemingly effortless manner in which they link a revolutionary idea to a traditional sentiment. It matters too that the overlay of images dramatizes an objective idea, and that what is remote and different is the fact that the Earth Tennyson invokes here "in tracts of fluent heat began":

*There rolls the deep where grew the tree.
O earth, what changes hast thou seen!
There where the long street roars hath
been
The stillness of the central sea.*

Maurice Riordan is at 6 Daniels Road, London SE15 3LR, UK. His most recent collection of poems is *Floods* (Faber & Faber).

Natural progression

Steven A. Benner

The seventeenth-century Enlightenment generated two very different approaches to thinking about the natural world. The first sought mathematical models to account for observations, beginning with the motion of the planets in the night sky. This approach generated 'physical science', a web of mathematical and physical models describing and predicting the behaviour of units of matter smaller than the atom, larger than the planets, and much in between.

The second approach emphasized collecting and classification. The archetypal practitioner was the English gentleman, who named plants, rocks and fossils on his estate, then in his empire. From this enterprise came 'natural history', a web of classification systems, non-reductive models and historical statements about the planet and the life it carries. The positions of continents came to be understood as outcomes of historical movements of plates on the globe, obeying physical laws but too complicated to be predicted by them. Life is also understood as the consequence of historical events, occurring within the context of darwinian theory, again consistent with physical laws but not predicted by them.

It is curious that these two very different activities receive the same name: 'science'. This has generated much confusion, and some rancour, with each group feeling itself to be the more 'relevant'. But the disjunction between the two traditions is (I believe) about to end, and in a largely unheralded way.

Many chemical biologists and biophysicists view the future of biology as a metamorphosis in which 'biological' phenomena will be replaced by their underlying 'physical chemical' components. This metamorphosis is ongoing, and very productive, but is unlikely to be the entire story. The surprise will come when biophysicists and chemical

biologists discover that they need to research the history of biomolecules if they are truly to understand the physical behaviours that they have worked so hard to characterize.

An early sign of the need for natural history in the physical sciences was the struggle to predict how proteins fold. This seemed to be a good area for applying the physical-science paradigm to biology, so physical chemists mounted a frontal assault, using huge computers to build physical models of proteins in water and making guesses about how atoms interact. The assault failed. The only way to make the computation even vaguely tractable, it turned out, required considerable abstraction of the physical model for the protein. The same physical theory that inspired the computation suggested that these abstractions must compromise the computation's value as a predictive tool.

Natural history offered an entirely different approach. Divergent evolution creates families of proteins that have descended from common ancestors. As proteins evolve from those ancestors, natural selection requires them to remain 'fit'. The principal prerequisite for fitness in a protein is a fold. So proteins diverging from a common ancestor generally conserve their folds. This means that during the evolution of protein sequences, mutations do not accumulate as they would if proteins were formless, functionless organic molecules. Instead, amino acids that are important to the fold suffer substitution differently from those that are not. A signal should lie in the pattern of protein-sequence divergence — the difference between how proteins have divergently evolved in their past, and how they would have evolved had they been formless, functionless molecules.

Natural history did not overcome the challenges obstructing the frontal assault of the physical scientists; it went around them. Today, the secondary and tertiary structure of proteins can reliably be predicted by exploiting the historical signal embedded in a set of protein sequences related by common ancestry. Since 1990, about 30 protein folds have been predicted using the history of protein families. In many cases, the prediction provided information about function as well as form.

Genomics is driving the use of history in physical science. Genomic-sequence databases contain historical information about genes in an easy-to-use form. It can be used to build evolutionary trees and reconstruct the sequences of ancestral genes and proteins. Through recombinant-DNA technology, ancient proteins from extinct organisms can be resurrected in the laboratory and studied

Molecular history

"The human genome, now no more than a set of chemical structures of the organic molecules involved in inheritance, will need natural history to give it value."

biochemically to test hypotheses about form and function. These techniques promise to deliver a comprehensive model for life, combining the physical and structural behaviour of biomolecules with two histories — the first told by palaeontological and geological records, the second by molecular sequences.

I believe that this historical model is as important for understanding the behaviour of biomolecules as the powerful instruments used for their physical characterization. Indeed, the human genome, now no more than a set of chemical structures of the organic molecules involved in inheritance, will need natural history to give it value.

Palaeogenomics models have already suggested how the global unification will look when it is complete. For example, a historical analysis of resurrected proteases ancestral to those regulating blood pressure indicated which animal models were appropriate for studying the pharmacology of pharmaceuticals targeted against these enzymes.

Combining natural history and physical science involves huge collaborations and faces many obstacles, not least of which is training. Compartmentalization of academic departments and funding bodies ensures that natural history is only rarely incorporated into biophysics curricula and that the converse is true for natural historians. But the potential rewards are immense. The first to reconcile the two traditions will be the first to glimpse the 'how' and 'why' of life from the molecule to the planet.

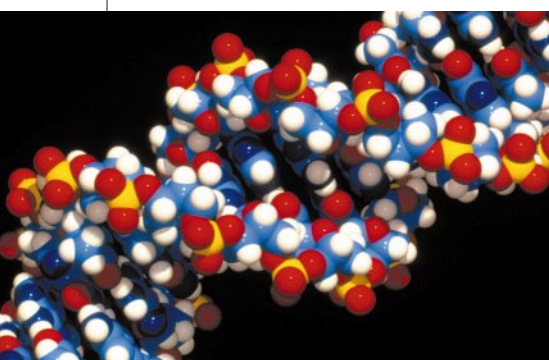
Steven A. Benner is in the Biochemistry Division of the Chemistry Department, University of Florida, PO Box 117200, Gainesville, Florida 32611, USA.

FURTHER READING

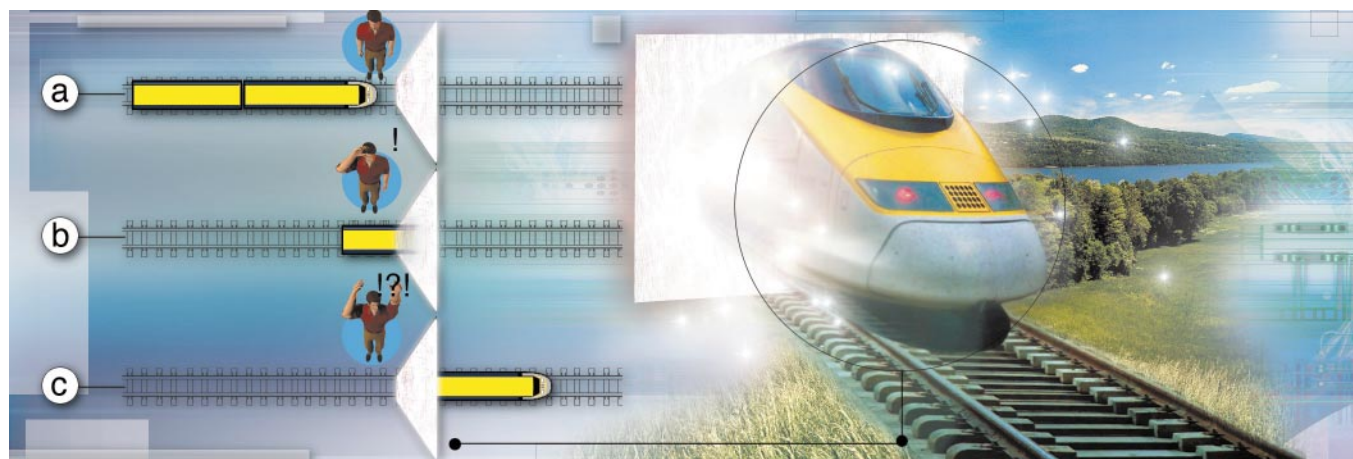
- Chandrasekharan, U. M. *et al. Science* **271**, 502–505 (1996).
- Jermann, T. M. *et al. Nature* **374**, 57–59 (1995).
- Behe, M. *Darwin's Black Box* (Free Press, New York, 1996).
- Benner, S. A. *et al. Chem. Rev.* **97**, 2725–2843 (1997).
- Benner, S. A. *et al. Adv. Enzyme Regul.* **38**, 155–180 (1998).
- Goh *et al. J. Mol. Biol.* **299**, 283–293 (2000).

WEBLINKS

- ▶ <http://www.stanford.edu/class/history133/Smoc/Unifying.html>
- ▶ <http://www.embl-heidelberg.de/predictprotein/predictprotein.html>



Interweaving the histories of biomolecules should deliver a comprehensive model for life.



Stopping light in its tracks

Eric A. Cornell

Using lasers and ultracold atoms, physicists have found a way to stop and start a pulse of light. This magic trick may one day be used to store data in a quantum computer.

Imagine you are standing beside a railway track, waiting for the next express train. Stretched across the tracks in front of you is a sheet of some strange, iridescent fabric, as thin as silk, which the train is about to rip apart (a, above). But, to your amazement, when the front of the train hits the gossamer fabric it does not immediately break through — instead it appears to be ‘soaked up’ into the material (b). In just a few seconds the entire train has vanished into the fabric. The roar of the train is replaced by an eerie silence. Then, as suddenly as it disappeared, the train emerges from the other side of the fabric (c), resuming its original length and speed, and roars off out of sight.

This trainspotting fantasy helps us to appreciate the most recent accomplishments (presented on page 490 of this issue¹) of Lene Hau and her colleagues. Two years ago, Hau’s group shot a pulse of laser light 3 microseconds in duration and about 1 kilometre in length into a specially prepared sample of ultracold sodium gas². The gas sample was about 0.2 mm in length and had the unusual property that the velocity of light within it was ten million times slower than in free space. When the leading edge of the pulse entered the sodium cloud, it immediately slowed to the unheard of speed of 30 m s^{-1} . At this leisurely pace, the light pulse took so long to cross the sample that, long before it emerged from the other side, the tail end of the light pulse vanished into the sample as well. Squeezed to within one ten-millionth of its original length, the pulse crept across the sample until finally it emerged, restored to its original length, and accelerated to its customary speed of $3 \times 10^8 \text{ m s}^{-1}$.

The key to slowing light is the presence of a second laser beam, the so-called ‘coupling’ pulse. Distinguishable from the propagating (or ‘probe’) pulse by its polarization, the coupling light delicately adjusts the internal energy levels of the atoms, suppressing their ability to absorb the probe light — in effect, a single absorption level is split into two levels that cancel each other out. This phenomenon is known as electromagnetically induced transparency³. At the same time, the ‘refractive index’ of the atomic cloud — in simple terms, how much it bends light — develops a steep dependence on the probe frequency. This in turn leads to a very slow ‘group velocity’ — the speed at which the envelope of light intensity moves through the sample.

Much of the excitement over the original experiment by Hau’s group focused on the value of the speed of light obtained. And in the past two years, several other groups have achieved similar results^{4–6}, using more conventional, room-temperature gases. In some cases, the speed of light was reduced to even slower than 30 m s^{-1} . Casual observers of the fracas were left to wonder, why bother with the ultracold atoms? But a more careful reading of the ‘hot-atom’ papers shows that they all lacked the eerie quality that so perturbed our hapless hobbyist by the train tracks. The maximum delay that the room-temperature experiments could impose on a light pulse was much less than the total duration of the pulse itself. So there was no chance that a significant fraction of the probe pulse (let alone the entire pulse) could be tucked away inside the sample of atoms. From a technological viewpoint, if we regard

the propagating light pulse as a ‘bit’ of optical information, then the hot-atom experiments displace each bit by only a small fraction of its width — not enough to turn a one into a zero, by any means.

The new results from Hau’s group¹ bring the sense of wonder back to slow-light experiments. In their latest caper, the group changes the rules of the game right in the middle of the observation. This time there is a magician standing beside the rail track. At the precise moment when the entire express train disappears into the fabric, the magician snaps her fingers. Click. And nothing happens. No train reappears. Finally, after a time long enough (had the magician not intervened) for the train to have gone another 200 km along the track, she snaps her fingers again and shazam! the train comes blasting out as if nothing had happened. Or, she can snap her fingers three times in quick succession and three separate carriages of the train appear.

What Hau and her colleagues actually do is cool a sample of sodium atoms to within one millionth of a degree above absolute zero. With the coupling-beam intensity tuned to provide a group velocity of 28 m s^{-1} , they inject a probe pulse 2 km long into the sample. Then, just as the pulse disappears into the sample, but before it reappears, the coupling beam is turned off. As a result, the speed of the probe beam goes to zero as well, and the pulse comes to a dead stop in the middle of the sample. The pulse of light can be kept on ice in the sample for up to a millisecond (a virtual eternity on the scale of the original 6-microsecond pulse duration). At any time in that millisecond interval,

the coupling beam can be turned back on, resulting in the probe beam promptly emerging from the sample and continuing along its way.

How can a photon be brought to a halt? Aren't these fundamental particles of light meant to travel, after all, at the speed of light? The key fact here is that, as the pulse of light penetrates into the dense region of the ultracold atomic cloud, it turns into a 'quantum coherence pattern' printed on the sodium atoms—the information in the light beam becomes stored within the quantum phase relationship within the internal atom states. In the final limit, when the pulse comes to a dead stop, all the photons have been 'imprinted' (absorbed in a fully reversible way) into the coherence pattern. Later, when the coupling light is turned back on, the information contained in the pattern is read out and converted back into propagating photons that accelerate to the conventional speed of light as they come to the edge of the atom sample.

While the probe pulse is in captivity, it is subject to the capricious powers of its jailers. By rapidly turning the group velocity up and down, Hau's group can let a small fraction of the pulse escape the sample, while keeping the remainder confined. In this way, they can chop up the incoming pulse into as many as three transmitted

pulses, with an arbitrary delay between each pulse. The authors refer to this as 'multiple read-out' of the stored pulse. Moreover, if the probe pulse is 'written in' to the material at a certain group velocity, and then is 'read out' at a higher group velocity, the transmitted pulse will have a shorter duration, and correspondingly higher intensity, than the injected pulse. It is as if the fantasy express train is now shorter and wider than the original.

Hau and her colleagues suggest that their newly demonstrated ability to control the flow of optical information may have technological relevance to quantum computing. Whether this will turn out to be true is not clear. But for now it hardly matters — trainspotting just doesn't get any more interesting than this. ■

Eric A. Cornell is at JILA, National Institute of Standards and Technology and the University of Colorado, Boulder, Colorado 80309-0440, USA.
e-mail: cornell@jila.colorado.edu

1. Liu, C., Dutton, Z., Behroozi, C. H. & Hau, L. V. *Nature* **409**, 490–493 (2001).
2. Hau, L. V., Harris, S. E., Dutton, Z. & Behroozi, C. H. *Nature* **397**, 594–598 (1999).
3. Harris, S. E. *Phys. Today* **50**, 36–42 (1997).
4. Kash, M. M. et al. *Phys. Rev. Lett.* **82**, 5229–5232 (1999).
5. Budker, D., Kimball, D. F., Rochester, S. M. & Yashchuk, V. V. *Phys. Rev. Lett.* **83**, 1767–1770 (1999).
6. Wang, L. J., Kuzmich, A. & Dogariu, A. *Nature* **406**, 277–279 (2000).

thirds of the 50-plus extrasolar planets discovered to date. The newest objects² are in orbit around a Sun-like star called HD168443, located about 123 light years away in the direction of the constellation Serpens.

Previous measurements showed that HD168443 had at least one planet³, but further observations were required to reveal what else might be lurking within the system. The latest data show that the original planet has a mass of at least 7.5 M_J and an orbital period of 58 days, placing it closer to its star than Mercury is to our Sun (about 0.3 astronomical units (AU) away from HD168443, where 1 AU is the Earth–Sun distance of 150 million kilometres). The second object has a mass of at least 17 M_J and orbits the star with a period of 4.8 years, at a distance of 3 AU, corresponding to the distance to the Sun of the asteroid belt in our Solar System. Both objects have eccentric (non-circular) orbits, as do essentially all the extrasolar planets orbiting at these distances.

Because the radial-velocity technique measures only the Doppler shift along the line of sight to the Earth, the actual masses could be much higher if their orbits are seen nearly face-on (rather than edge-on). But this is unlikely, so the masses are probably only about 30% above the minimum values. The absence of a visible wobble in the position of HD168443 seen in astrometric data from the Hipparcos satellite² limits the mass of the outer companion to less than 42 M_J . Together, HD168443's known companions have at least 25 times the mass of the planets in our Solar System. We're not in Kansas, any more, Toto.

Figure 1 shows how the new system compares to the bestiary of extrasolar planets and brown dwarfs previously discovered around Sun-like stars. There is no obvious break in the mass distribution because nature has managed to populate nearly the entire range of possibilities that have been searched to date, leaving no niche unoccupied (much like the behaviour of terrestrial organisms). The two objects orbiting HD168443 fall uncomfortably close to, or in the middle of, the provisional dividing line between gas giants and brown dwarfs, thought to lie between roughly 10 M_J and 30 M_J . The radial-velocity technique favours the detection of heavyweights, so objects with these masses should be very easy to detect. Yet the fact that very few objects have so far been found in this range suggested that there might be a gap between the masses of the most massive gas giants and the least massive brown dwarfs, however rare the latter might be. But HD168443 throws a monkey wrench into these feeble attempts to classify what we've found so far.

Beyond simple classification, there is the looming question of how a system like

Extrasolar planets

Giant giants or dwarf dwarfs?

Alan P. Boss

An object at least 17 times the size of Jupiter, discovered orbiting a Sun-like star, has astronomers scratching their heads. Is it a giant planet or a failed star?

Planet hunters keep discovering yet more weird objects around nearby

Sun-like stars that resemble nothing found in our own Solar System. The latest discoveries, announced at a meeting earlier this month*, include a three-body system that appears to be composed of a central star orbited by two objects with masses no smaller than 7.5 and 17 times that of Jupiter, the largest planet in the Solar System. The larger of these two behemoths is massive enough to be termed a brown dwarf star. Brown dwarfs are failed stars that are large enough (greater than about 13 Jupiter masses, M_J) to burn some deuterium, but not massive enough to ignite the steady nuclear reactions that make stars shine for billions of years. The discovery of this giant threatens to help overturn the tentative distinctions astronomers have tried to draw between gas giant planets, such as Jupiter,

and brown dwarf stars. But reclassifying these beasts may be an almost trivial task compared with figuring out how they formed in the first place.

A worldwide race has been underway for the past five years¹ to complete the first census of Jupiter-like planets around Sun-like stars in our local neighbourhood of the Milky Way. The preferred method for detecting these objects is the 'radial-velocity technique', in which telescopes equipped with high-precision spectrographs measure the Doppler shifts of starlight caused by the star orbiting a common centre of mass with its unseen companions. Jupiter, for example, causes the Sun to wobble in space by about 12.5 metres per second during its 12-year orbit.

The latest announcements come from the incredibly successful team led by Geoff Marcy (University of California, Berkeley) and Paul Butler (Carnegie Institution of Washington), who have found roughly two-

* 197th Meeting of the American Astronomical Society, San Diego, California, USA, 7–11 January 2001.

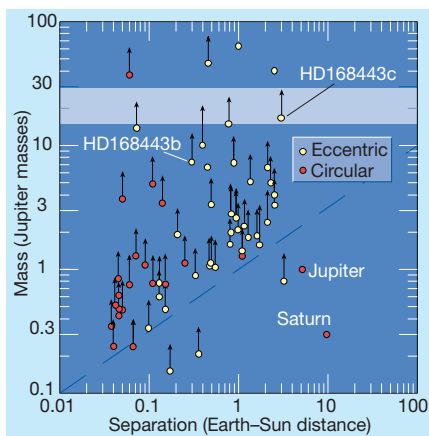


Figure 1 Pattern of discoveries for gas giant planets and brown dwarf stars found orbiting nearby stars similar to our Sun. In most cases, only the minimum mass of the companion is known, although this value is likely to be close to the true mass. The star HD168443 has now been found to have two massive companions (HD168443b and HD168443c) that blur the distinction between giant planets and brown dwarf stars. In particular, the mass of HD168443c falls into the observed gap (from about 15 M_J to 30 M_J) between gas giants and brown dwarfs. The radial-velocity technique preferentially detects massive, close-in planets, as indicated by the detection bias along the dashed line.

HD168443 might be created. Could it have the same history as a multiple star system? These are thought to start with the collapse of a dense molecular cloud core to form a flattened protostellar disk, followed by fragmentation into multiple protostars. Our understanding of this process is still hazy, but it seems an unlikely way to give birth to a tight stellar system in which the mass of the primary star is 50 to 100 times more massive than its companions. This is because subsequent growth by gas accretion from the disk and an infalling envelope of gas would add mass preferentially to the smaller protostars, leading to more equal stellar masses⁴. Furthermore, fragmentation calculations do not favour the formation of low-mass protostars in stable, planet-like orbits around a single much more massive protostar. When multiple protostars form by fragmentation, they begin life in unstable configurations, which are then subject to gravitational scattering during close mutual encounters.

We are therefore left with trying to explain the formation of the HD168443 system by the process of planetary formation which unfolds well after the central protostar has gained the bulk of its mass. There are two competing mechanisms for the formation of gas giants in a protoplanetary disk: core accretion and disk instability. The currently favoured theory, core accretion, was developed to explain the formation of Jupiter and Saturn through the formation of

a rock and ice core, followed by the accretion of gas. Core accretion could account for the formation of planets with masses up to about 5 M_J (refs 5–7), but whether it could produce objects as massive as the companions of HD168443 within the lifetime⁸ of a typical protoplanetary disk (a few million years) remains to be seen. Disk instability⁹, on the other hand, is quite capable of rapidly forming massive protoplanets, as large as 17 M_J , with change to spare.

Theorists have their work cut out to explain the formation of HD168443's unexpected companions. The planet hunters, meanwhile, are discovering bizarre solar systems at an alarming rate.

Microbiology

Gastrogenomics

Jonathan A. Eisen

The genome of an *Escherichia coli* strain that is emerging as a severe threat to human health has been sequenced. Comparing it with that of a harmless strain suggests why some forms of this bacterium cause disease.

Humans and *Escherichia coli* normally live happily together: *E. coli* is a beneficial bacterium commonly found in the human gastrointestinal system. But it also exists in harmful forms. One of the most harmful of these, called O157:H7, was first linked to human disease in 1983 (ref. 1), when it was shown to have been the cause of two outbreaks of an unusual and severe gastrointestinal ailment in the United States the previous year. The number of documented human illnesses and deaths caused by O157:H7 strains has since increased steadily worldwide, and these strains are now considered to be both emerging pathogens and major threats to public health². Studies of O157:H7 receive a boost from a paper on page 529 of this issue³, in which Perna and colleagues describe and analyse the genome sequence of one strain of *E. coli* O157:H7.

Outbreaks of O157:H7 infections in humans have been traced primarily to infected cattle⁴, which are the source of contamination of ground beef, milk and — indirectly, through fertilizer — many fruit and vegetable products. Fortunately, proper cooking or pasteurization can prevent human infection by contaminated food. But infections can also be transmitted by sewage-contaminated water and person-to-person contact.

Genetically, the pathogenicity (ability to cause damage) and virulence (degree of pathogenicity) of O157:H7 strains depend on several factors. For example, these strains possess genes encoding the so-called shiga toxin, as well as small, circular DNA molecules that encode 'virulence factors'. And

Alan P. Boss is in the Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road NW, Washington DC 20015-1305, USA.

e-mail: boss@dtm.ciw.edu

1. Mayor, M. & Queloz, D. *Nature* **378**, 355–359 (1995).
2. Marcy, G. W. et al. *Astrophys. J.* (submitted). See <http://exoplanets.org>
3. Marcy, G. W., Butler, R. P., Vogt, S. S., Fischer, D. & Liu, M. C. *Astrophys. J.* **520**, 239–247 (1999).
4. Bate, M. R. *Mon. Not. R. Astron. Soc.* **314**, 33–53 (2000).
5. Bodenheimer, P., Hubickyj, O. & Lissauer, J. J. *Icarus* **143**, 2–14 (2000).
6. Kley, W. *Mon. Not. R. Astron. Soc.* **313**, L47–L51 (2000).
7. Nelson, R. P., Papaloizou, J. C. B., Masset, F. & Kley, W. *Mon. Not. R. Astron. Soc.* **318**, 18–36 (2000).
8. Briceño, C. et al. *Science* **291**, 93–96 (2001).
9. Boss, A. P. *Astrophys. J.* **536**, L101–L104 (2000).

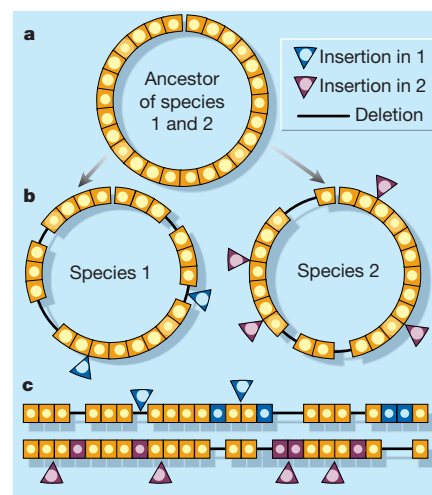


Figure 1 Model of the generation of genome islands specific to just one of a pair of related strains or species. a, The ancestral circular genome. b, Modern genomes — the results of differential insertion and deletion of genomic segments in each lineage. c, Genomes 1 and 2 linearized and aligned. 'Islands' within each genome are highlighted on alignment (blue for species 1 and purple for species 2). Note that islands in one genome are generated both by insertion of genetic segments into that genome and deletion of segments from the other genome.

O157:H7 has at least one pathogenicity island — a section of chromosomal DNA containing many genes that contribute to pathogenicity⁵. Evolutionary studies have shown that all O157:H7 strains are closely related, and that they share a common ancestry with pathogenic O55:H7 strains⁶

(the numbers and letters refer to the different types of variable antigen molecules produced by these bacteria). Many virulence factors in O157:H7 strains were acquired from other strains or species by a process known as horizontal gene transfer^{5,7}. So, what can analysis of the whole genome add to our knowledge of this pathogen?

Perna *et al.*³ have used the now standard 'whole-genome shotgun' approach to determine the genome sequence of the O157:H7 strain EDL933, which was isolated from ground beef linked to the 1982 outbreak. Their analysis of this genome, and in particular their comparison of this genome with that of a non-pathogenic laboratory *E. coli* strain, K-12 MG1655, is revealing.

The two genomes have a large amount of DNA — about 4.1 million base pairs (megabases) each — that was clearly derived from a common ancestor. This 'backbone' DNA is arranged similarly in the two strains: the two genomes can be lined up side by side along their lengths, except at one point, where part of the O157:H7 genome is reversed.

Although this conserved arrangement and inversion are not surprising for closely related strains⁸, one feature is rather unusual. Scattered roughly evenly within each genome's backbone are hundreds of sections of DNA that are unique to one or the other strain. Sections found only in O157:H7 — 'O-islands' — total 1.34 megabases and 1,387 genes. K-islands, which are unique to the non-pathogenic *E. coli* strain, add up to 0.53 megabases and 528 genes. It remains to be seen which of these differences contribute to the virulence and pathogenicity of O157:H7. The O-islands include many known and predicted pathogenicity genes — for example, some of them may encode toxins, or factors needed to make the adhesive filaments (fimbriae) that help the bacterium to stick to the lining of the gut. But it is difficult to predict gene function accurately, and many of the O-islands might have no connection with pathogenicity.

Differences between the DNA backbones may also be important. Although most of the backbone differences do not result in changes in protein sequence, many do: about 75% of the backbone-encoded proteins differ by at least one amino acid between the two strains. A more thorough analysis of these patterns will help in determining which differences are the result of natural selection and which are merely neutral changes.

Interestingly, the patterns of variation within each genome differ between the coding and non-coding strands of backbone genes. Perna *et al.* suggest that this may result from transcription-coupled repair of oxidative damage in DNA. This process was originally discovered in *E. coli*: as the coding strand is copied into RNA (transcribed),

DNA damage in that strand is mended at a higher rate than normal⁹. Confirmation of whether this has caused the strand bias described here will require analysis of the genome sequence of another related species¹⁰.

The authors also suggest that much of the DNA in the O-islands and K-islands was acquired by horizontal gene transfer. One of their lines of evidence is that many of the islands contain sequences related to those of bacterial viruses and other vectors that carry genes from one species to another. Another possibility is that the islands were present in the common ancestor of the two strains, and then lost in one lineage (Fig. 1).

Analyses of the genomes of related species will also help to answer this question. But if horizontal gene transfer has occurred, then the fact that so many genes have been transferred to O157:H7 supports the suggestion¹¹ that the continuing emergence of O157:H7 as a pathogen results from its ability to undergo rapid genetic change. This suggestion was made because a high proportion of O157:H7 strains have defects in genes involved in repairing DNA mismatches¹¹. This tends to lead to higher rates of both mutation and acquisition of DNA from other strains. Many other pathogenic bacteria also have mismatch-repair defects¹², but so too do many non-pathogenic *E. coli* strains¹³, and the existence of so many K-islands suggests that gene transfer is also common in non-pathogenic strains. Moreover, O157:H7 has an apparently normal long-term rate of sequence change¹⁴.

Perna *et al.*'s work³ emphasizes the power of comparing genomes from closely related strains or species — something that is becoming possible for more and more taxa. Such comparisons allow the detection and analysis of genetic processes that occur on relatively short timescales. They have led to discoveries such as the possible occurrence of transcription-coupled repair of oxidative damage, reported here³, and the finding that inversions that are symmetrical around the start point of replication of a bacterial chromosome are common in bacterial genome evolution⁸.

Many of these insights depend on knowing details such as gene location and orientation, and the absence of genes that are present in related species. This emphasizes the importance of having complete or nearly complete genome sequences. (The sequencing of the O157:H7 genome is nearly complete; only two gaps remain.) We should view with scepticism press releases announcing the completion of a genome sequence in a day¹⁵ — they refer simply to the completion of the initial sequencing part of a project, but many gaps (frequently hundreds) always remain. Closing those gaps is difficult but essential.

There is still much to learn about the



100 YEARS AGO

An interesting description of the ravages of white ants, or termites, in Rhodesia is furnished by the Rev. A. Leboeuf to the *Zambesi Mission Record* for January. The special interest of the contribution centres in the account of the damage done to property by white ants in Rhodesia, which seems to be even greater than in India. It is no uncommon thing, says the writer, for the colonist, on returning from his day's labour, to find the coat he left hanging on a nail of his cottage wall and the books on the table absolutely destroyed by these tiny marauders. Nor is this all. "On awaking next morning," writes Mr. Leboeuf, "you are astonished to see in the dim light a cone-shaped object rising from the brick floor a short distance from your bed, with two holes on the top like the crater of a miniature volcano. Upon closer examination you discover that the holes have just the size and shape of the inside of your boots, which you incautiously left on the brick floor the night before. They have given form and proportion to an ant heap, and nothing is left of them except the nails, eyelets and, maybe, part of the heels."

From *Nature* 24 January 1901.

50 YEARS AGO

May I support Prof. Alan Boyden's plea that parthenogenesis should not be classified as asexual reproduction? The habit of doing so presumably arose because the text-book definition of sexual reproduction excludes anything which does not involve fusion of gametes; but such a reliance on a definition instead of on the facts implies a degree of philosophical realism which has no proper place in science. Biology, unlike logic and mathematics, has to take the world as it finds it. Definitions are descriptions of concepts or phenomena, sometimes of arbitrary stages in a series, and when they are made short they inevitably become inaccurate... No satisfactory definition of 'species' has ever been made. One of the latest, that of Mayr, is quite inadequate; apart from being formally incorrect in stating that a species is a group of populations, which it certainly is not, his definition would, since it contains the terms 'interbreeding', exclude *Amoeba proteus* and all creatures without cross-fertilization. Similar difficulties arise with the words cell, reproduction, tissue, skeleton, parasite and many others.

From *Nature* 27 January 1951.

biology and history of *E. coli* O157:H7, but the sequence and analysis presented by Perna *et al.*³ will be an excellent starting point. One of the great things about genomics is that the data it provides allow nearly any group, anywhere in the world, to start seeking answers to outstanding questions immediately. ■

Jonathan A. Eisen is at the Institute for Genomic Research, 9712 Medical Center Drive, Rockville, Maryland 20850, USA.

e-mail: jeisen@tigr.org

1. Riley, L. W. *et al.* *New Engl. J. Med.* **308**, 681–685 (1983).
2. Mead, P. S. *et al.* *Emerg. Infect. Dis.* **5**, 607–625 (1999).
3. Perna, N. T. *et al.* *Nature* **409**, 529–533 (2001).
4. Wells, J. G. *et al.* *J. Clin. Microbiol.* **29**, 985–989 (1991).

5. Reid, S. D., Herbelin, C. J., Bumbaugh, A. C., Selander, R. K. & Whittam, T. S. *Nature* **406**, 64–67 (2000).
6. Whittam, T. S. *et al.* *Infect. Immun.* **61**, 1619–1629 (1993).
7. Pupo, G. M., Karaolis, D. K., Lan, R. & Reeves, P. R. *Infect. Immun.* **65**, 2685–2692 (1997).
8. Eisen, J. A., Heidelberg, J. F., White, O. & Salzberg, S. L. *Genome Biol.* **1**, research0011.1–0011.9 (2000).
9. Mellon, I., Spivak, G. & Hanawalt, P. C. *Cell* **51**, 241–249 (1987).
10. Pilar-Francino, M., Chao, L., Riley, M. A. & Ochman, H. *Science* **272**, 107–109 (1996).
11. LeClerc, J. E., Li, B., Payne, W. L. & Cebula, T. A. *Science* **274**, 1208–1211 (1996).
12. Eisen, J. A. *Nucleic Acids Res.* **26**, 4291–4300 (1998).
13. Matic, I. *et al.* *Science* **277**, 1833–1834 (1997).
14. Whittam, T. S., Reid, S. D. & Selander, R. K. *Emerg. Infect. Dis.* **4**, 615–617 (1998).
15. http://news.bbc.co.uk/1/hi/english/sci/tech/newsid_743000/743329.stm

Carbon cycle

The age of river carbon

Wolfgang Ludwig

The organic carbon that runs into the oceans from rivers could be hundreds or thousands of years old. If so, aspects of our understanding of the global carbon cycle will have to change.

Dissolved organic carbon in the oceans is one of the biggest reservoirs in the global carbon cycle — it is comparable in size to all of the carbon in terrestrial plants, or to all of that in the form of CO₂ in the atmosphere. The input of terrestrial organic carbon from rivers, the main source of most constituents of sea water¹, could fill this marine reservoir in just a few thousands of years² — which, according to radiocarbon dating³, is also about the average age of marine organic carbon. But although there should be lots of terrestrial-derived carbon in the ocean, geochemical studies indicate that there seems to be very little. So what happens to the riverine carbon once it enters the oceans?

On page 497 of this issue⁴ Raymond and Bauer address the question by presenting new data on the average ages of organic matter in rivers. There may in fact be much more terrestrial carbon in the oceans than we thought, but we cannot see it. To understand this, one has to know how geochemists usually distinguish between organic carbon from terrestrial and marine sources.

Almost all of the organic carbon on Earth is created through photosynthesis, whether on land or in water. But on land the process leaves characteristic fingerprints, which mean that this carbon should be traceable after it has entered the oceans. Many land plants synthesize certain compounds, such as lignin or tannin, which are absent in marine phytoplankton. In principle, then, detecting these biomarkers in the sea can reveal if carbon had a terrestrial origin. The other widely applied method involves measuring the ratio between the two stable carbon isotopes, ¹³C and ¹²C, in the bulk organic matter. Most land plants produce carbon that is



Figure 1 The Amazon delta, seen from space. Raymond and Bauer⁴ find that, contrary to general belief, most of the organic carbon entering the Atlantic Ocean from the Amazon is not 'fresh' but has aged and been degraded.

more depleted in the heavy carbon isotope (¹³C) than carbon produced by marine phytoplankton, leading to higher isotopic ratios in marine than in terrestrial carbon.

When going into details, however, things can be more complicated. Biomarker concentrations in different plant materials are variable, reducing their potential for the accurate assessment of sources. And stable carbon-isotope ratios can sometimes give ambiguous results because there are exceptions to the general trend⁵. A broader problem arises when comparing relatively old organic material in the sea with freshly synthesized material of terrestrial and marine plants, because chemical and biological degradation may have altered the composition of the fresher material and make it difficult to trace its origin.

This is where Raymond and Bauer's study⁴ comes in. They analysed organic matter in the Amazon (Fig. 1) and three North American rivers by radiocarbon dating, and

found that it was up to several thousand years old. This contrasts starkly with the general belief that most of the organic carbon in rivers should be relatively 'fresh'⁶. The particulate organic carbon (that is, the fraction retained on a filter) was especially old, the dissolved carbon being younger on average. But even the dissolved fraction revealed relatively old ages in some samples, and long-term laboratory experiments with samples from one of the rivers showed that the slow bacterial oxidation of part of this organic matter can markedly increase the average age of the carbon that remained in solution.

From these results Raymond and Bauer conclude that ageing and degradation of terrestrial organic matter in river basins and coastal zones may significantly alter its structure, distribution and quantity before it reaches the open oceans. The implication is that we simply have not been able to distinguish it from marine-generated carbon. Future studies need to test whether the old radiocarbon ages found by the authors can be confirmed in rivers from other parts of the world, however, and whether ageing can wipe out the fingerprints of terrestrial carbon. Furthermore, old average ages of riverine carbon can also be explained by the presence of small amounts of 'fossil' organic carbon⁷ that originated in ancient sedimentary rocks. This phenomenon may be restricted to rivers draining these kind of rocks, and that has to be tested as well.

But it remains possible that dissolved organic carbon in the sea is mainly composed of old marine carbon, and that the terrestrial component is not hidden but largely absent. What if that is the case? A process would then be needed that rapidly removes the terrestrial carbon after it enters the oceans. It is generally accepted that part of this carbon is buried in the coastal zones where sedimentation rates are high. Biological activity is also high in these waters, and some carbon may be released as CO₂ to the atmosphere through biological oxidation. But the two processes do not seem to be powerful enough to remove all riverine carbon⁵, and an additional mechanism for its rapid withdrawal from the sea would remain on the most-wanted list in the offices of many carbon-cycle researchers.

Nevertheless, at the least the results of Raymond and Bauer⁴ remind us that the organic matter that runs from rivers into the sea is not necessarily identical to the organic matter of the plants and soils upstream in the river catchments. Certainly, other processes taking place on the floodplains or in the river channels, such as slowed transport caused by cycles of deposition and movement downstream, may also have to be considered to better understand the composition and fluxes of river carbon⁸. ■

Wolfgang Ludwig is at the Centre de Formation et

de Recherche sur l'Environnement Marin (CEFREM), CNRS/Université de Perpignan, 52 Avenue de Villeneuve, 66860 Perpignan Cedex, France.
e-mail: ludwig@univ-perp.fr

- Mackenzie, F. T. & Garrels, R. M. *Am. J. Sci.* **264**, 507–525 (1966).
- Smith, S. V. & Hollibaugh, G. T. *Rev. Geophys.* **31**, 75–89 (1993).

- Williams, P. M. & Druffel, E. R. M. *Nature* **330**, 246–248 (1987).
- Raymond, P. A. & Bauer, J. E. *Nature* **409**, 497–500 (2001).
- Hedges, J. I., Keil, R. G. & Benner, R. *Org. Geochem.* **27**, 195–212 (1997).
- Ludwig, W., Probst, J.-L. & Kempe, S. *Glob. Biogeochem. Cycles* **10**, 23–41 (1996).
- Kao, S. & Liu, K. *Limnol. Oceanogr.* **41**, 1749–1757 (1996).
- McClain, M. E., Richey, J. E. & Brandes, J. A. *Glob. Biogeochem. Cycles* **11**, 295–311 (1997).

Biomechanics

Walking on other planets

Alberto E. Minetti

A nineteenth-century equation used for building model ships allows us to compare the motion of animals of different sizes and gaits. It may also give us an idea of how we would move on different planets.

What do a gibbon swinging by its arms from tree to tree, a sailing steamship, and a walking human have in common? The answer lies in a simple concept introduced in the nineteenth century by nautical engineer William Froude to help him to produce model ships that maintained the same propulsion dynamics as full-size vessels. His 'Froude number' is a dimensionless variable that was brought into biology by D'Arcy Thompson¹ and popularized by McNeill Alexander^{2,3} in the study of the energetics and mechanics of animal locomotion. This number allows one to compare the motion of species with different numbers of legs and gaits, and to investigate the effects of different body sizes on the mechanics of movement. As suggested by various earlier papers^{4–6}, and borne out by new work described by Cavagna and colleagues in the *Journal of Physiology*⁷, the Froude number is also a reliable rule-of-thumb for predicting walking speeds on planets with gravities different from that of the Earth.

In walking humans or swinging gibbons, changes in the vertical position of the body's centre of mass during movement affect the gravitational potential energy of the body, and are accompanied by opposite changes in the kinetic energy needed to drive movement. The result is a pendulum-like mode of movement that saves mechanical energy. Roughly the same principles apply to ships, but here the potential energy is related to the size of the wave generated by the ship.

One of the theories underlying studies of movement — dynamic similarity³ — basically states that geometrically similar bodies that rely on pendulum-like mechanics of movement will have similar gait dynamics if the Froude number remains the same. This number (Fr) is given by $Fr = v^2/(g \times l)$, where v is the speed of movement (in $m\ s^{-1}$), g is acceleration due to gravity (in $m\ s^{-2}$) and l is a characteristic length (such as leg length, in metres). The Froude number is directly proportional to

the ratio between the kinetic energy and the gravitational potential energy needed during movement.

Dynamic similarity implies that, for example, despite differences in body size and

number of legs, humans and quadrupeds change from walk to run or trot at the same Froude number, close to 0.5 (ref. 3). Even gibbons⁸ change from merely swinging to swinging with aerial phases at Froude numbers ranging from 0.3 to 0.6. Humans of short height, such as children^{9,10}, patients suffering from early-onset growth-hormone deficiency¹¹ and pygmies¹², optimally walk at speeds that correspond dynamically to that for adults, that is, with a Froude number of 0.25. Here, walking speed is optimal when the recovery of energy by the body — by exchanging potential and kinetic energies — is maximal⁷. What all of this means is that, within the same gravitational environment (such as on the Earth), the smaller the body, the lower the 'equivalent' speed of movement, which is proportional to the square root of the leg length.

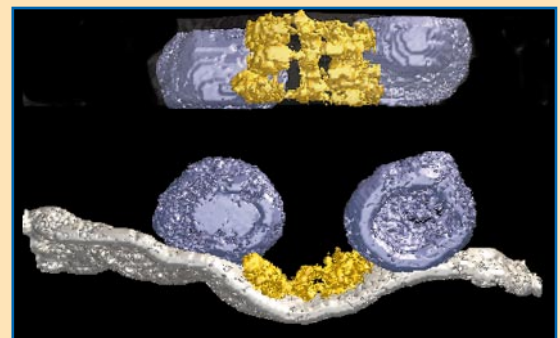
But the power of the Froude number for predicting equivalent walking speeds is not confined to the Earth. Within the same species and for a given body size, dynamic

Neurobiology

Activity at the active zone

Presynaptic active zones are the specialized sites from which nerves release neurotransmitter. They are characterized by the presence of calcium channels, synaptic vesicles containing neurotransmitter, and aggregates of protein that make up what is known as 'active-zone material'. This material was described by S. L. Palay almost 50 years ago (*J. Biophys. Biochem. Cytol.* **2**, 193–202; 1956), but its precise structure and function have remained an enigma. Things may be about to change, however, thanks to the publication elsewhere in this issue of stunning three-dimensional reconstructions of the frog neuromuscular junction (*Nature* **409**, 479–484; 2001).

To study this junction between nerve and muscle, M. L. Harlow and colleagues used a technique known as electron microscope tomography, which gives greater spatial resolution than conventional transmission electron microscopy. They describe three distinct elongate structures — which they call ribs, beams and pegs — that make up the first 15 nanometres



of active-zone material as one looks into the neuron. The top image shows this view, with the active-zone material in yellow and the synaptic vesicles in blue.

Beams — which make up the vertical band in the centre of this image — are arranged along the long axis of the elongate active zone, and form the backbone of the structure. Ribs lie perpendicular to this long axis, and link the central backbone to the double row of synaptic vesicles that straddle it. Pegs (not apparent in these images) tether each rib to the plasma membrane in a pattern that matches the distribution of molecules that are thought to be calcium channels. The lower image is a side-on view, with the presynaptic plasma

membrane in grey. The structure is perhaps best appreciated in the movies published on *Nature's* web site (<http://www.nature.com/nature/>).

As a whole, the active-zone material appears to provide a scaffold by which synaptic vesicles are localized to the specialized presynaptic membrane. But the proteins that make up the structure probably have more specific functions. The intracellular domains of calcium channels may form all or part of the pegs, and ribs may be composed of the host of proteins that together mediate docking and fusion of vesicles with the presynaptic membrane. We eagerly await the formal identification of the rib, beam and peg proteins. **Lesley Anson**

M. L. HARLOW ET AL.

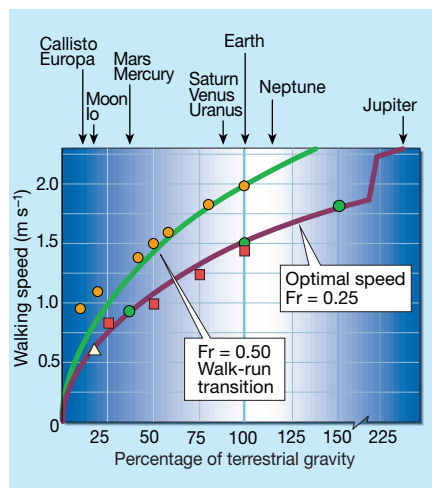


Figure 1 Walking speed as a function of gravity. The curves represent the predictions of the theory of dynamic similarity. These predictions were made using the Froude-number equation, $Fr = v^2/(g \times l)$, where v is the velocity in m s^{-1} , g is the acceleration due to gravity in m s^{-2} , l is 0.92 m (the average leg length for adult male humans), and Fr is 0.25 (which sets the optimal walking speed on Earth^{9–12}) or 0.5 (at which humans move spontaneously from a walk to a run on Earth³). Squares and orange circles refer to measurements of optimal walking speeds⁴ and walk-to-run-transition speeds⁵, respectively, in simulated low-gravity conditions. The triangle represents a previous estimate of reduced walking speed on the Moon¹³. Green circles represent recent measurements of optimal speed obtained using parabolic flight to simulate low- and high-gravity conditions^{6,7}.

similarity also predicts that the lower the gravity, the slower the equivalent walking speed, which depends on the square root of the gravity ratio — the gravity of the planet in question divided by that of the Earth. On Earth, the optimal walking speed and walk-to-run-transition speed for an adult man of average height are, respectively, about 1.5 and 2.0 m s^{-1} . Figure 1 shows the relationships between walking speed and gravity at Froude numbers of 0.25 (which determines the optimal walking speed for humans) and 0.5 (which determines the walk-to-run-transition speed). One can use the Froude number to predict that the corresponding speeds on a stellar body with 16% of the Earth's gravity, such as the Moon, will be about 40% (the square root of 16%) of those on Earth — that is, about 0.6 and 0.8 m s^{-1} , respectively.

These values were also predicted on the basis of a different approach about 35 years ago¹³, and the impossibility of walking at terrestrial speeds on the Moon has been clear to see in debriefings from the Apollo missions¹⁴. Moreover, experimental evidence has confirmed the predictive power of the Froude number in low gravity. Different low-gravity conditions have been simulated

in the lab by applying a nearly constant upward force to the waist of subjects walking on a motorized treadmill^{4,5}. As predicted, the optimal walking speed⁴ and the spontaneous transition between walking and running⁵ occurred at Froude numbers close to 0.25 and 0.5, respectively. The slight discrepancy observed at low gravity (Fig. 1) is likely to be caused by the approximate nature of the simulation — limbs were allowed to swing normally, as if still affected by terrestrial gravity.

Cavagna *et al.*⁷ now go further, to look at the mechanics of human walking in both low and high gravity. The low-gravity conditions correspond to those on Mars, being 40% of the Earth value. The high-gravity value was 150% of that on Earth. The authors used given portions of the parabolic trajectory of an aeroplane equipped with platforms that were sensitive to forces in all directions. In this way, Cavagna *et al.* measured the displacement of the body's centre of mass during a couple of walking steps at different constant gravities and walking speeds. From these measurements, they calculated the recovery of energy. They also recorded the speeds at which energy recovery was optimal (the optimal walking speeds). These speeds closely match those predicted when the Froude number is 0.25 (ref. 15), although the range of optimal speeds in high gravity is quite broad.

Indeed, Cavagna *et al.* conclude that increased gravity increases the range of walking speeds. This is again predicted by the Froude number, as seen in Fig. 1, where the vertical distance between the curves increases with increasing gravity. Here, the vertical distance represents the range of speeds within two dynamically different conditions.

Plant biology

Floral quartets

Günter Theißen and Heinz Saedler

Goethe was right when he proposed that flowers are modified leaves. It seems that four genes involved in plant development must be expressed together to turn leaves into floral organs.

What controls the difference between a plant's floral organs and its leaves? Over 200 years ago Johann Wolfgang von Goethe proposed that the different parts of a plant result from 'metamorphosis' (meaning transformation) of a basic organ, the 'ideal leaf'. But if floral organs are just modified leaves, what are the modifiers? On page 525 of this issue¹ Honma and Goto provide the answer in molecular terms.

A typical flower consists of four different types of organ arranged in four whorls (Fig. 1, overleaf). There are leaf-like sepals in the outermost whorl; showy petals in the second

Bipedal and quadrupedal walking — unlike swinging — remain an approximation of ideal pendulum dynamics. Despite this, the Froude number, and the underlying theory of dynamic similarity, has so far proved a handy rule-of-thumb for predicting equivalent walking speeds as a function of body size or gravity. It is not yet known to what extent faster walking speeds can be maintained in high gravity for more than the few steps studied by Cavagna *et al.*⁷. But it is nonetheless impressive that an idea introduced in the nineteenth century to produce model ships can be used in the twenty-first century to predict how humans would walk on other planets.

Alberto E. Minetti is in the Department of Exercise and Sport Science, Manchester Metropolitan University, Hassall Road, Alsager ST7 2HL, UK. e-mail: a.e.minetti@mmu.ac.uk

1. Thompson, D. W. *On Growth and Form* (Cambridge Univ. Press, 1961).
2. Alexander, R. M. *Nature* **261**, 129–130 (1976).
3. Alexander, R. M. *Physiol. Rev.* **69**, 1199–1227 (1989).
4. Griffin, T. M., Tolani, N. A. & Kram, R. *J. Appl. Physiol.* **86**, 383–390 (1999).
5. Kram, R., Domingo, A. & Ferris, D. J. *Exp. Biol.* **200**, 821–826 (1997).
6. Cavagna, G., Willems, P. A. & Heglund, N. C. *Nature* **393**, 636 (1998).
7. Cavagna, G., Willems, P. A. & Heglund, N. C. *J. Physiol. (Lond.)* **528**, 657–668 (2000).
8. Cheng, Y. H., Bertram, J. E. & Lee, D. V. *Am. J. Phys. Anthropol.* **113**, 201–216 (2000).
9. Cavagna, G. P., Franzetti, P. & Fuchimoto, T. *J. Physiol. (Lond.)* **343**, 232–339 (1983).
10. Dejaeger, D., Willems, P. A. & Heglund, N. C. *Pflügers Arch. Eur. J. Physiol.* **441**, 538–543 (2001).
11. Minetti, A. E., Ardigo, L. P., Saibene, F., Ferrero, S. & Sartorio, A. *Eur. J. Endocr.* **142**, 35–41 (2000).
12. Minetti, A. E. *et al. Eur. J. Appl. Physiol.* **68**, 285–290 (1994).
13. Margaria, R. & Cavagna, G. *Aerospace Med.* **35**, 1140–1146 (1964).
14. Minetti, A. E. *Proc. R. Soc. Lond. B* **265**, 1227–1235 (1998).
15. Minetti, A. E. *Acta Astron.* (in the press).

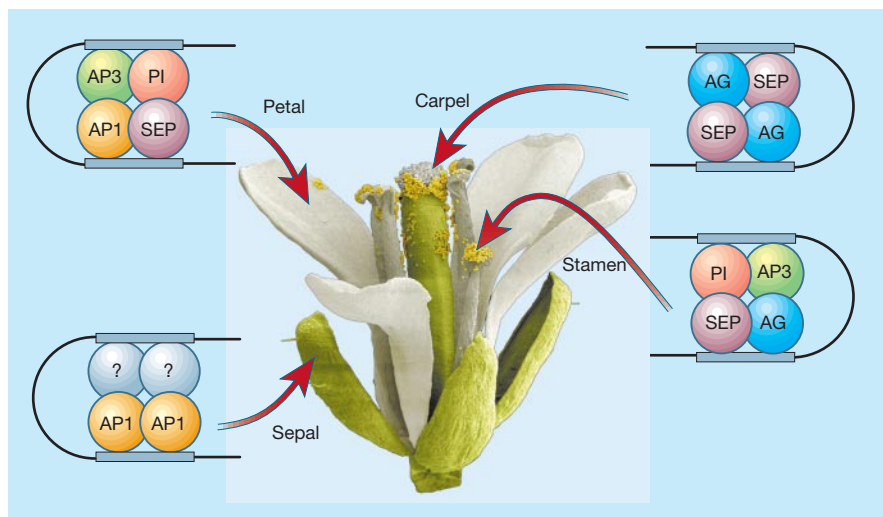


Figure 1 Flower structure and the 'quartet model'⁶ of floral organ specification in *Arabidopsis*. According to this model, the identity of the different floral organs — sepals, petals, stamens and carpels — is determined by four combinations of floral homeotic proteins known as MADS-box proteins^{1,5,8}. The protein quartets, which are transcription factors, may operate by binding to the promoter regions of target genes, which they activate or repress as appropriate for the development of the different floral organs. According to the model, two dimers of each tetramer recognize two different DNA sites (termed CArG-boxes, shown here in grey) on the same strand of DNA, which are brought into close proximity by DNA bending. The exact structures of the multimeric complexes of MADS-box proteins controlling the identity of flower organs are still hypothetical; question marks denote components whose identity is especially uncertain. Proteins: AG, AGAMOUS; AP1, APETALA1; AP3, APETALA3; PI, PISTILLATA; SEP, SEPALLATA.

mutants have petals in the third and sepals in the fourth whorl^{2,3}. Organ identity in the other whorls is unchanged.

Following genetic analyses of these mutants, simple models were developed to explain how the different floral organs adopt their unique identities during development^{2,3}. The most widely known is the ABC model, in which combinatorial interactions between three classes of floral homeotic genes are affected in the respective mutants. So these genes are termed class A, class B and class C genes, with A specifying sepals, A + B petals, B + C stamens and C carpels³. (Class D genes, specifying ovules, were later added to the ABC model, but they are not considered further here.) The *Arabidopsis* class A genes are *APETALA1* (AP1) and *APETALA2* (AP2); the class B genes are *APETALA3* (AP3) and *PISTILLATA* (PI); and the only class C gene is *AGAMOUS* (AG). Apart from AP2, these genes are all members of the MADS-box family^{2–4} which encode transcription factors — proteins that recognize specific DNA motifs of other genes and influence their transcription.

Combined loss-of-function of class A, B and C genes results in a transformation of all floral organs into leaves, corroborating Goethe's view that leaves are a developmental 'ground state'. But expression of the ABC genes throughout a plant does not transform leaves into floral organs, showing that the ABC genes, though necessary, are not sufficient to superimpose floral

organ identity on a leaf developmental programme^{1,5}. So, what other mysterious factors are required to specify floral organ identity?

A telling clue came just a few months ago with the report⁵ that all flower organs resembled sepals in mutants where three other MADS-box genes — *SEPALLATA1* (*SEP1*), *SEP2* and *SEP3* — had lost their function. Together with gene expression data, this finding suggested that the *SEP* genes represent another class of floral homeotic genes, termed class E genes⁶, which, together with the class B and C genes, is required for the specification of organ identity in the petal (A + B + E), stamen (B + C + E) and carpel (C + E)^{5,6}.

But by what mechanism do these different genes interact? MADS-box proteins bind DNA as dimers, but attempts to explain the interaction between class A, B and C genes by MADS-box protein heterodimerization have failed⁷. However, there was a twist to the tale with the demonstration⁸ that some homeotic MADS-box proteins from *Antirrhinum* form multimeric DNA-binding complexes. Honma and Goto¹ now report that *Arabidopsis* proteins also bind to DNA as multimeric complexes containing the class B proteins AP3 and PI, the class E protein SEP3, and either AP1 (a class A protein) or AG (a class C protein). These are the exact combinations of proteins required to specify petal (A + B + E) or stamen (B + C + E) identity, respectively. The ability of MADS-box proteins to form multimeric

complexes may therefore provide the molecular basis for the combinatorial interaction of the floral homeotic genes.

So are the *SEP* proteins the mysterious factors that, along with the ABC genes, are required to specify floral organ identity? The answer is clearly 'yes'. In genetically engineered plants that express B + (A or E) genes throughout development, Honma and Goto found that leaves are transformed into petal-like organs, and that ubiquitous expression of class B, C and E genes transforms leaves into stamen-like organs.

With these findings^{1,5}, improved successors of the ABC model can be developed. An example is the 'quartet model' which directly links floral organ identity to the action of four different tetrameric transcription factor complexes composed of MADS-box proteins (Fig. 1; for details, see ref. 6). Goals for the future will be to define the exact structures of these transcription-factor complexes inside the living plant cell, to identify the target genes whose transcription is regulated by the binding of the complexes, and to explain the gene specificity of that binding.

The results of Honma and Goto¹ also promise progress in answering one of the enduring puzzles of botany: the evolutionary origin of the flower⁹. The identity of floral organs is totally dependent on the activity of the MADS-type floral homeotic genes, so gene duplication and diversification within the MADS-box gene family must have been key processes in flower evolution⁴.

The MADS-type floral homeotic genes can be separated into three different gene clades (sets of genes that share a last common ancestor not shared with any of the other MADS-box genes)^{4,10}. From phylogenetic reconstructions it seems that these clades — including the class B, class C + D and class A + E genes, respectively — were established within a relatively short period of time well before the origin of the flowering plants. But once established, these gene clades have changed very little^{4,10}.

Does this unusual pattern of gene evolution reflect the 'invention' of heterotetramer formation and subsequent coevolution of the constituents? Were novel specificities in DNA-binding⁸ and in the regulation of target genes, generated by the establishment of tetramer formation, required for the specification of more sophisticated reproductive organs, such as flowers or cones? Such questions can now be answered by studying the phylogeny, function and interaction of MADS-box proteins in various groups, including non-flowering plants^{4,9,10}.

It was prescient of Goethe to recognize floral organs as modified leaves. But is there even more to it? In his novel *Die Wahlverwandtschaften* (*The Elective Affinities*),

Goethe described the rapid establishment of new ways of sexual reproduction when a married couple (a 'dimer') decided to live together with another man and woman. By inventing this quartet, did Goethe even anticipate the molecular basis of floral-organ specification?

Günter Theißen and Heinz Saedler are in the Department of Molecular Plant Genetics, Max Planck Institute for Plant Breeding Research, Carl-von-Linné-Weg 10, D-50829 Köln, Germany. e-mails: theissen@mpiz-koeln.mpg.de
saedler@mpiz-koeln.mpg.de

1. Honma, T. & Goto, K. *Nature* **409**, 525–529 (2001).
2. Schwarz-Sommer, Z., Huijser, P., Nacken, W., Saedler, H. & Sommer, H. *Science* **250**, 931–936 (1990).
3. Coen, E. S. & Meyerowitz, E. M. *Nature* **353**, 31–37 (1991).
4. Theißen, G. *et al.* *Plant Mol. Biol.* **42**, 115–149 (2000).
5. Pelaz, S., Ditta, G. S., Baumann, E., Wisman, E. & Yanofsky, M. F. *Nature* **405**, 200–203 (2000).
6. Theißen, G. *Curr. Opin. Plant Biol.* **4**, 75–85 (2001).
7. Riechmann, J. L., Krizek, B. A. & Meyerowitz, E. M. *Proc. Natl Acad. Sci. USA* **93**, 4793–4798 (1996).
8. Egea-Cortines, M., Saedler, H. & Sommer, H. *EMBO J.* **18**, 5370–5379 (1999).
9. Crepet, W. L. *Proc. Natl Acad. Sci. USA* **97**, 12939–12941 (2000).
10. Purugganan, M. D. *Bioessays* **20**, 700–711 (1998).

Quantum physics

Watching an atom tunnel

Ali Yazdani

Physicists have managed to watch individual hydrogen atoms move on metal surfaces at very low temperatures — in defiance of classical physics.

Quantum theory, which has just celebrated its 100th birthday, allows particles to break the rules of classical physics. Although, in a classical sense, a particle may not have sufficient energy to cross a given barrier, quantum theory says that there is still a definite probability that it can penetrate the barrier in a process known as quantum tunnelling. This process is at the heart of many phenomena — from the formation of chemical bonds to what distinguishes a metal from an insulator. Writing in *Physical Review Letters*, Lauhon and Ho¹ report that they have tracked and visualized the quantum tunnelling of individual atoms for the first time.

By using a scanning tunnelling microscope (STM), the researchers were able to monitor the motion of individual hydrogen atoms on a metal surface. They found that the atoms remain mobile down to temperatures as low as 9 K. Classically, thermal diffusion or motion is expected to fade away as

the temperature is lowered. But the constant movement of hydrogen implies that there is a quantum effect that allows the atoms to tunnel along the metal's surface. Quantum tunnelling of atoms at low temperatures has been inferred from experiments on relatively large groups of atoms, but never before has quantum motion been observed so directly — one atom at a time.

In condensed-matter physics, quantum tunnelling of atoms is believed to play a key role in phenomena such as the diffusion of impurities in solids and the properties of glasses at low temperatures. An atom typically 'rests' in an energy well (Fig. 1a). It can tunnel to another well if it is light enough and if the energy barrier between the wells is sufficiently small. Because hydrogen is so light, it is particularly open to the possibility of quantum tunnelling². On the surfaces of metals, the constant movement of hydrogen has been reported down to low temperatures^{3,4}. But whether this

diffusion arises from classical thermal motion or from quantum tunnelling is unclear. Some of the uncertainty can be attributed to the fact that previous experiments measured the average behaviour of a group of atoms, and so could not resolve the role of surface defects in the tunnelling process. A localized probe, such as the tip of an STM, sidesteps these complications.

The STM uses the principle of electron tunnelling to give us a close-up view of atoms at surfaces⁵. A stream of electrons tunnels between a sharp metallic tip and the surface under investigation as the tip skims across the surface. A feedback circuit keeps the flow of electrons constant by adjusting the distance between the tip and the surface. The recorded trajectory of the tip is then used to form an image.

The presence of a hydrogen atom on the surface is seen as a dip or 'hole' on an STM image (Fig. 1b). This is because an STM image of an isolated atom on a surface depends on how much the atom modifies the surface's electronic structure. Isolated hydrogen atoms provide very little attraction for the conduction electrons of a metal surface and, as Lauhon and Ho¹ show, they appear only as small dips in the STM images of copper at low temperatures (9 K). An identical image is obtained for deuterium, hydrogen's heavier isotope, because the two are electronically equivalent. But when hydrogen and deuterium are chemically bound to the surface, the STM can distinguish between them because their different masses cause them to vibrate at different frequencies⁶.

The rate at which atoms move across a surface can be measured by making an STM 'movie' of the sample. Alternatively, the tip can lock on to an atom and keep constant track of its location. Using both techniques, Lauhon and Ho measured the rate of diffusion of hydrogen and deuterium across a copper surface at different temperatures. From their observations, they ruled out the effect of the STM tip on their measurements. At higher temperatures, they saw both atoms obeying a diffusion law that is consistent with classical motion caused by thermal excitation. But as the surface temperature was lowered, the behaviour of hydrogen and deuterium began to differ dramatically. Deuterium follows classical physics and slows down to a halt. But hydrogen shows a very weak temperature-independent motion, which starts at about 65 K and is still present at 9 K, the lowest temperature used in the experiment. Because hydrogen is much lighter than deuterium, this movement provides strong evidence that quantum tunnelling is at work.

A particle's ability to tunnel is also affected by its immediate environment. For example, lattice vibrations (phonons) and conduction electrons scattering from the atom as

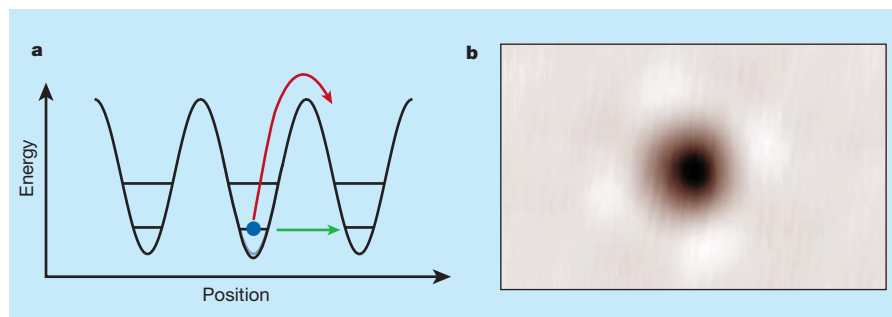


Figure 1 Hydrogen atoms on a copper surface can quantum tunnel at low temperatures. a, A hydrogen atom in an energy well created by the copper surface. The red path shows the energy needed for the atom to diffuse into the next energy well under classical physics. The green path indicates that the atom needs little or no energy to tunnel by quantum mechanical means into the adjacent well. b, An image of a hydrogen atom (dark feature in the centre) on a copper surface taken by scanning tunnelling microscopy.

it tunnels can hinder or promote the tunnelling action. The interaction of a tunnelling atom with its environment dissipates energy and can disrupt the tunnelling process, reduce it to a mere hop, or stop it altogether⁷. Detailed analysis of the STM-measured diffusion rate for hydrogen by Lauhon and Ho identifies different temperature regimes in which phonon- or electron-scattering predominates. Most intriguing is their observation that at the lower temperatures in the experiment, the tunnelling rate increases as the surface is cooled. This classically impossible behaviour shows that hydrogen tunnelling improves over longer periods of time as the surface gets colder. Perhaps reducing the temperature by a further factor of 10 or 100 will reveal a new regime in which hydrogen atoms can eventually tunnel over greater distances.

Nowadays, the focus of much tunnelling research is to determine the degree to which this quantum concept can be extended to the macroscopic world⁸. As the size of the tunnelling object is increased, the effect of the environment becomes much more

important. The situation in Lauhon and Ho's experiments is still microscopic, yet it provides a remarkably clean setting to get at the heart of environmental effects on the tunnelling process. The STM can also be used to position atoms carefully so as to design different tunnelling hurdles and to examine how the tunnelling process can be controlled. Perhaps learning how to manipulate the outcome of the tunnelling events at the nanometre scale would have some useful applications in making devices on this scale.

Ali Yazdani is in the Department of Physics and the Frederick Seitz Materials Research Laboratory, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA.
e-mail: ayazdani@uiuc.edu

1. Lauhon, L. J. & Ho, W. *Phys. Rev. Lett.* **85**, 4566–4569 (2000).
2. Völkl, J. & Alefeld, G. in *Hydrogen in Metals I* (eds Alefeld, G. & Völkl, J.) Ch. 1 (Springer, New York, 1978).
3. DiFoggie, R. & Gomer, R. *Phys. Rev. B* **25**, 3490–3511 (1982).
4. Daniels, E. A. & Gomer, R. *Surf. Sci.* **336**, 245–261 (1995).
5. Yazdani, A. & Lieber, C. M. *Nature* **401**, 227–230 (1999).
6. Stipe, B. C., Rezaei, M. A. & Ho, W. *Science* **280**, 1732–1735 (1998).
7. Leggett, A. J. *et al. Rev. Mod. Phys.* **59**, 1–85 (1987).
8. Blatter, G. *Nature* **406**, 25–26 (2000).

Neurobiology

A new code for axons

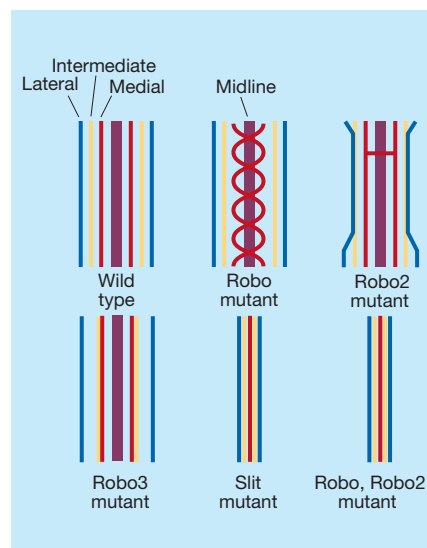
Guy Tear

Neurons extend axons, which must grow along specific pathways during development. The discovery of particular patterns of expression of so-called Robo proteins provides clues to how axons find their way.

During development, neurons send out projections called axons, which must grow along an appropriate route — selecting from a multitude of possibilities — to reach their target. In organisms that are bilaterally symmetrical, many axons must also cross the central line (the midline) of the central nervous system, forming connections called commissures, which join the two sides of the body. Other axons, meanwhile, remain on one side. Four papers in *Cell* and *Neuron*^{1–4} now reveal the surprising result that, in the fruitfly *Drosophila melanogaster*, a protein called Slit, expressed at the midline, has a say both in controlling midline crossing and in directing an axon's subsequent choice of route. These activities require the previously known Roundabout (Robo) protein, found on the surface of neurons, and two newly discovered members of the same family.

In the central nervous system (CNS) of insects and vertebrates, neurons initially project axons towards specialized cells at the midline. The axons then turn into a longitudinal pathway, either before or after crossing the midline. The specialized midline cells secrete an axonal repellent — Slit — which binds to Robo⁵. Certain axons are attracted

to midline cells and express Robo on their surface only after crossing the midline. This stops them from crossing back over. Meanwhile, axons that remain on one side of the CNS express Robo continuously. In other words, by signalling through Robo, Slit hangs up a 'do not enter' sign, stopping axons from coming into the midline.



In the absence of Robo, many axons freely cross and recross the midline⁶ (Fig. 1). But when Slit function is experimentally eliminated, all axons remain at the midline. This suggests that Slit may normally signal through other receptors in addition to Robo, providing a 'get out' signal. As now shown by Simpson *et al.*¹ and Rajagopalan *et al.*², one of the new Robo proteins, Robo2, is required for this signal. Robo2 is first expressed on all longitudinal axons in a pattern comparable to that of Robo. Eliminating the function of both Robo and Robo2 has the same effect as eliminating Slit — axons are unable to leave the midline (Fig. 1).

Having made its initial decision about crossing the midline, each axon turns into one of about 20 different longitudinal pathways, which are positioned at various distances either side of the midline. Axons in the various longitudinal pathways express certain surface markers that identify them. For example, the cell-adhesion molecule Fasciclin II is expressed in three major axon pathways either side of the midline.

The existence of such restricted expression patterns gave rise to the 'labelled pathways' hypothesis, which states that axons use these adhesive surface cues to decide which pathway to enter⁷. What has not been clear, however, is how the axons are directed the appropriate distance from the midline before they recognize the surface cues. Simpson *et al.*³ and Rajagopalan *et al.*⁴ now suggest that the two new Robo proteins — Robo2 and Robo3 — provide a combinatorial code that sets the sensitivity of an axon to the midline signal Slit, and hence the distance the axon travels from the midline (Fig. 2).

At the time that axons choose their longitudinal route, Robo2 is expressed only on those axons that extend within the most lateral third of the longitudinal pathways — that is, those pathways furthest from the

Figure 1 Guiding axons across the midline of the central nervous system. The figure shows the trajectories taken by Fasciclin-II-expressing axons in wild-type fruitflies and in fruitflies with mutations in Slit or one of its binding partners, the Robo proteins. These axons extend in three pathways at differing distances from the midline; the medial pathway is the closest and the lateral pathway the most distant. As shown in two new papers^{1,2}, Robo and Robo2 cooperate to regulate the crossing of the midline by axons. Robo, Robo2 and Robo3 together regulate the lateral position of the longitudinal pathways.

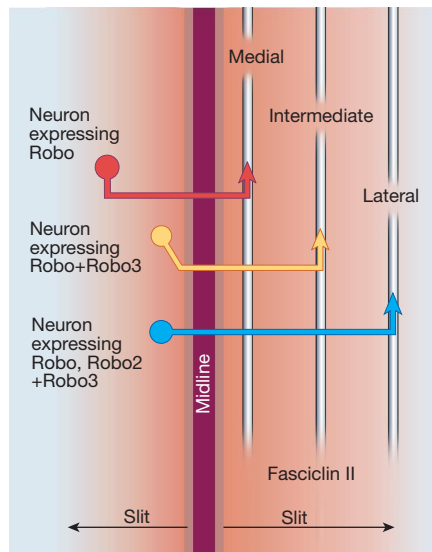


Figure 2 How axons choose their pathway either side of the midline. This model is suggested by two new papers^{3,4}. The Slit protein, secreted by midline cells, diffuses across the central nervous system. Axons that express Robo proteins from the start do not cross the midline but head a certain distance towards the midline according to the combination of Robo proteins that they express (not shown). Axons that do not initially express Robo and Robo2 cross the midline. The distance they then travel across the central nervous system depends on the combination of Robo proteins that they come to express (shown here). Axons expressing Robo and Robo3 extend into the intermediate zone, whereas axons expressing Robo, Robo2 and Robo3 are more strongly repelled from the midline and enter a lateral zone. Once in the correct zone, axons use specific surface markers (such as Fasciclin II) to select the appropriate pathway.

midline. The expression of Robo3 likewise begins at this time, and is restricted to axons that extend within the lateral two-thirds of the longitudinal pathways. Robo, by contrast, is expressed by all axons in the longitudinal pathways.

The final, overlapping pattern of expression of Robo, Robo2 and Robo3 sets up a series of zones within the longitudinal pathways (Fig. 2). Each zone includes one of the Fasciclin-II-expressing pathways. The medial pathway — that nearest to the midline — is defined by the expression of Robo, the intermediate pathway by Robo and Robo3, and the lateral pathway by all three Robo proteins. Experimentally changing the combination of Robo molecules expressed by the axons drives them to extend within different longitudinal domains^{3,4}. For example, eliminating Robo3 causes axons to extend into a more medial domain, resulting in a fusion of the intermediate Fasciclin-II-expressing pathway with the medial pathway. Similarly, the forced expression of Robo2 in neurons expressing

only Robo causes their axons to extend more laterally.

In other words, the specific combination of Robo proteins expressed by an axon — its 'Robo code' — determines the approximate longitudinal region it reaches. The final choice of route is then mediated by local attractive interactions between other surface markers. This combination of long-range repulsion (in this case mediated by Slit, produced at the midline, and the Robo proteins, found on axons) and local cues in the target area has also been suggested as a mechanism for the topographical mapping of axons in the brain's tectum^{8,9}.

This model suggests that, although Slit is expressed by midline cells, the protein must be distributed across the CNS. Although a gradient of Slit in the CNS has not yet been detected, it surely exists. Cells must be able to detect Slit at some distance from the midline, as the migration of muscle precursor cells at the lateral edge of the CNS is known also to be regulated by Slit⁵. The Robo molecules bind Slit, albeit with differing efficiencies, and switches in pathway choice based on changes in the Robo code produce the predicted discrete axonal movements towards or away from the midline, supporting the idea of a graded signal.

Do all the Robo proteins use the same intracellular signalling pathway to control axon extension in response to Slit? Robo2 and Robo3 lack two of Robo's four conserved intracellular domains, which bind to the intracellular signalling molecules Enabled

and Abl¹⁰. So the newly identified Robo molecules are likely to transmit a qualitatively different intracellular signal, although what this is remains unknown. Moreover, Simpson *et al.*¹ show that the Robo molecules can form homodimers and heterodimers, which is likely to affect the signalling output.

How is the graded Slit signal generated? How do the Robo molecules transmit their signals into the axon? And is this type of combinatorial code also used in vertebrates to direct and position neurons? The number of species in which Slit-to-Robo signalling has been found is increasing rapidly^{11–13}, so it probably won't be long before we find the answers to these questions.

Guy Tear is at the MRC Centre for Developmental Neurobiology, 4th Floor, New Hunt's House, Guy's Hospital Campus, King's College, London SE1 1UL, UK.

e-mail: guy.tear@kcl.ac.uk

1. Simpson, J. H., Kidd, T., Bland, K. S. & Goodman, C. S. *Neuron* **28**, 753–766 (2000).
2. Rajagopalan, S., Nicolas, E., Vivancos, V., Berger, J. & Dickson, B. J. *Neuron* **22**, 767–777 (2000).
3. Simpson, J. H., Bland, K. S., Fetter, R. D. & Goodman, C. S. *Cell* **103**, 1019–1032 (2000).
4. Rajagopalan, S., Vivancos, V., Nicolas, E. & Dickson, B. J. *Cell* **103**, 1033–1045 (2000).
5. Kidd, T., Bland, K. S. & Goodman, C. S. *Cell* **96**, 785–794 (1999).
6. Kidd, T. *et al.* *Cell* **92**, 205–215 (1998).
7. Goodman, C. *et al.* *Science* **225**, 1271–1279 (1984).
8. Brown, A. *et al.* *Cell* **102**, 77–88 (2000).
9. Feldheim, D. A. *et al.* *Neuron* **25**, 563–574 (2000).
10. Bashaw, G. J., Kidd, T., Murray, D., Pawson, T. & Goodman, C. S. *Cell* **101**, 703–715 (2000).
11. Sundaresan, V. *et al.* *Mol. Cell Neurosci.* **11**, 29–35 (1998).
12. Brose, K. *et al.* *Cell* **96**, 795–806 (1999).
13. Erskine, L. *et al.* *J. Neurosci.* **20**, 4975–4982 (2000).

Developmental biology

Sperm and mammalian polarity

Roger A. Pedersen

The location at which a sperm enters a mouse egg during fertilization helps to determine the layout of the embryo. This supports the view that mammalian embryos acquire their pattern early in development.

In many animal species, the point of entry of a sperm into an egg during fertilization determines the orientation in which the resulting one-celled embryo will divide to produce two cells. It also sets the orientation of one of the three embryonic 'axes', which form later in development and mark the three dimensions of the embryo: front to back, head to tail, and left to right. Until now, mammals were not thought to be included among these species. But Piotrowska and Zernicka-Goetz¹ change this by discovering a role for sperm in the formation of an early embryonic axis in mouse embryos (page 517 of this issue). This finding represents a leap forward in our knowledge of how the mammalian embryo acquires its body pattern. It also suggests that mammals might share other

features of axis formation with species, such as frogs, for which we have a better understanding of the effects of fertilization on embryonic organization.

This discovery comes at a time of changing views about what mammalian embryos accomplish in their first week of life, before they attach to the uterus. Until recently, this early (preimplantation) period was thought to be important mainly for the formation of nutrient-supplying and protective outer tissues, which surround the fetus during life in the uterus but are discarded at birth². But these outer tissues — which are known as 'extraembryonic' because they contribute no cellular descendants to the fetus — may actually be important in establishing the orientation and pattern of the mammalian body³. For example, the organ-

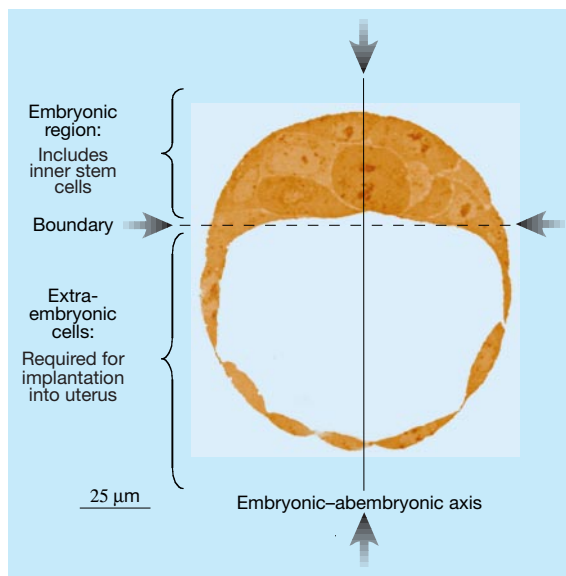


Figure 1 The 'preimplantation' mouse embryo at 3.5 days old. The figure shows the boundary between the embryonic region (which includes the inner stem cells that form the fetus) and the extraembryonic cells required for implantation of the embryo into the uterus. Piotrowska and Zernicka-Goetz¹ show that the position at which the sperm enters the mouse egg during fertilization predicts this boundary. The embryonic–abembryonic axis is perpendicular to the boundary. Photograph of embryo is reproduced with permission from ref. 11.

ization of one of the extraembryonic tissues is important for the future head-to-tail orientation of the body. Moreover, Zernicka-Goetz's group has previously shown that the polarity of the preimplantation embryo predicts the organization of this particular extraembryonic tissue⁴.

Further axes are needed to orient the body from back to front and from left to right. In frogs, sperm determine the back-to-front axis. Studies of the African frog *Xenopus laevis* have shown that fertilization sparks off a 30° rotation of the outermost layer of the egg (its cortex) relative to the egg's deeper contents (its cytoplasm)⁵. This rotation (away from the point of sperm entry) results in the alignment of an array of cytoskeletal components — the cortical microtubules — in the yolky half of the egg. These microtubules act as tracks, along which components of the so-called Wnt signalling pathway are transported towards the prospective back (dorsal midline) of the embryo⁶. The result is dorsal enrichment of the transcription factor β -catenin⁷, which switches on genes specific to the dorsal midline. The sperm's entry point also determines the plane of the first division of *Xenopus* embryos, but this appears to be accomplished through a different mechanism⁸.

The importance of the sperm's entry position is not unique to frogs. It sets the first cleavage plane and/or an embryonic axis in a wide range of species, including ascidians (sea squirts), molluscs and roundworms. Where studied, these animals have in common cortical or cytoplasmic reorganization, as well as periodic production of waves of calcium ions, related to the site of sperm entry⁹. But how similar are mice to these species?

Piotrowska and Zernicka-Goetz¹ have found that the sperm's point of entry into a mouse egg determines the plane of the first division of the embryo. It also influ-

ences the timing of the second cell division — the cell in the two-cell embryo that inherits the sperm entry point usually divides before the other cell. This might provide the early-dividing cell with a developmental advantage, such as preferential specialization to form one of the three cell lineages that develop during the pre-implantation period — two types of extraembryonic cell, and the inner stem cells, which form the fetus.

Moreover, the authors discovered that the sperm entry position marks the boundary between the extraembryonic cells needed for implantation of the embryo into the uterus, and the inner stem cells. This boundary defines the so-called 'embryonic–abembryonic' axis, which aligns perpendicular to the boundary (Fig. 1).

Piotrowska and Zernicka-Goetz have concentrated on the preimplantation period of mouse embryonic development. So their work tells us nothing about the later effects of sperm entry position on the development of the three body axes. This is a difficult issue to tackle: the mammalian embryo is very small at first (the size of a pinpoint), and it becomes inaccessible to study at just about the time of overt axis formation, because it embeds itself into the wall of the uterus. But we could speculate that, in forming the embryonic–abembryonic and the head-to-tail axes, the mouse embryo establishes the orientation of its body pattern before it implants into the uterus. This would be remarkable.

Further studies will be needed to determine whether or not the sperm entry site in mice also establishes the dorsal midline, as it does in frogs. In any event, finding a role for sperm in their embryonic patterning does bring mice closer to frogs and other non-mammalian species, and raises the question of whether similar molecular mechanisms underlie the effects of sperm on patterning

Daedalus

David Jones

David Jones, author of the Daedalus column, is indisposed.

in all vertebrates. Calcium waves starting from the site of sperm entry, and modest cytoplasmic movements, do occur in mouse embryos¹⁰, as they do in frogs. But little is known about the orientation and function of the cytoskeleton during the period after fertilization in mice, much less whether there is any enrichment of Wnt-pathway components in relation to the sperm entry position. The mouse embryo is peculiar even among mammals in having no conspicuous sperm-influenced cytoskeletal organization. So, if the mouse sperm does exert its organizational role through such a mechanism, its effects are more subtle than in other species.

Finally, there is a sobering implication of the new findings¹. During human infertility treatments, sperm are injected directly into human eggs. Might this be defining an embryonic axis, or even the future body pattern, of the child? With this possibility in mind, it becomes urgent to find out exactly how sperm affect mouse patterning. Although mouse embryos are more difficult to study than frogs, they are continuing to provide answers to questions that are clearly relevant to humans.

Roger A. Pedersen is in the Reproductive Genetics Unit, Department of Obstetrics, Gynecology and Reproductive Sciences, University of California, San Francisco, California 94143, USA.
e-mail: pedersen@cgl.ucsf.edu

1. Piotrowska, K. & Zernicka-Goetz, M. *Nature* **409**, 517–521 (2001).
2. Gardner, R. L. *Curr. Top. Dev. Biol.* **39**, 35–71 (1998).
3. Beddington, R. S. & Robertson, E. J. *Cell* **96**, 195–209 (1999).
4. Weber, R. J., Pedersen, R. A., Wianny, F., Evans, M. J. & Zernicka-Goetz, M. *Development* **126**, 5591–5598 (1999).
5. De Robertis, E. M., Larrain, J., Oelgeschläger, M. & Wessely, O. *Nature Rev. Genet.* **1**, 171–181 (2000).
6. Miller, J. R. et al. *J. Cell Biol.* **146**, 427–437 (1999).
7. Larabell, C. A. et al. *J. Cell Biol.* **136**, 1123–1136 (1997).
8. Danilchik, M. V. & Black, S. D. *Dev. Biol.* **128**, 58–64 (1988).
9. Roegiers, F., McDougall, A. & Sardet, C. *Development* **121**, 3457–3466 (1995).
10. Deguchi, R., Shirakawa, H., Oda, S., Mohri, T. & Miyazaki, S. *Dev. Biol.* **218**, 299–313 (2000).
11. Pierce, G. B. *Am. J. Pathol.* **113**, 117–124 (1983).

Correction

In Heike Langenberg's meeting report, "Meteorology: Oscillating opinion" (*Nature* **408**, 924–925; 2000), the "new index" for the North Atlantic Oscillation, attributed to P. Yiou, refers to a poster by V. C. Slonosky and P. Yiou. The index was developed by Victoria Slonosky, Phil Jones and Trevor Davies at the University of East Anglia. Further details can be found in Slonosky, V. C., Jones, P. D. & Davies, T. D. *Int. J. Clim.* **20**, 1874–1897 (2000).

Watching fights raises fish hormone levels

Cichlid fish wrestling for dominance induce an androgen surge in male spectators.

Social interactions among the animals in a group affect their subsequent behaviour, manifesting as dominance hierarchies or territoriality, for example, and meaning that behaviour is adjusted to social context¹. These interactions are thought to be modulated by androgens, allowing the agonistic motivation of individuals to adjust to changes in their social environment; androgen production is itself determined by sexual status and by social contacts among conspecifics². Here we show that the adaptive value of socially modulated androgen levels in male cichlid fish can be extended to bystanders who watch but who do not participate in social confrontation.

The 'challenge hypothesis' proposes that an individual's androgen production responds to their social interactions with other group members². We investigated whether this hypothesis can be extended to bystanders by monitoring the effect of watching fights between conspecifics on the androgen levels of adult male cichlid fish, *Oreochromis mossambicus*. Males of this species defend mating territories in leks during the breeding season.

Fish of comparable size and age and raised in similar conditions were placed in social isolation for 7 days (phase 1) to minimize the influence of previous social experience on androgen levels. Bystander males were allowed to watch two isolated conspecific male neighbours through a one-way mirror for a period of 3 days (phase 2). After this time, the opaque partition keeping the two neighbours apart was removed at 9 a.m. to allow a one-hour fight for dominance to take place. The fight was watched by the bystander through the one-way mirror (phase 3).

We collected urine at different stages from bystanders for radioimmunoassay to determine the concentrations of testosterone and 11-ketotestosterone, the principal fish androgens³. Samples were taken 2 hours before the start of phase 3 to act as a reference for androgen levels in the presence of two conspecific males; at 30 minutes after the end of the fight; at 2 hours after the end of the fight; and at 6 hours after the end of the fight. Urine from controls (who witnessed no fight because the opaque partition was kept in place, preventing interaction between the bystander's two neighbours) was also collected at the time points corresponding to these stages.

If androgens are produced in response to a social interaction among conspecifics, androgen concentrations in urine from

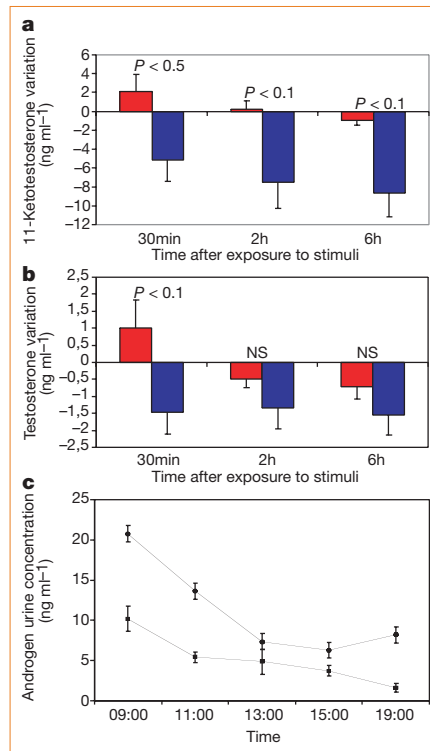


Figure 1 Social interactions modulate androgen levels of bystanders. **a**, **b**, Variation of bystander (red) total (free + sulphate + glucuronide) urinary levels of 11-ketotestosterone (**a**) and testosterone (**b**) after witnessing a conspecific fight relative to the pre-test reference values (concentration before watching the fight minus the concentration after watching the fight). Blue, controls. **c**, Daily variation in testosterone (bottom curve) and 11-ketotestosterone (top curve) in males after 7 days of social isolation. Sample size (n) was 18 fish, apart from the 11-h sample point, for which $n=17$, and the 11-ketotestosterone 19-h sample point, for which $n=13$. Urine was collected into 1.5-ml Eppendorf tubes by applying gentle pressure to the ventral area of the fish flanks. Sample processing radioimmunoassay characteristics have been described elsewhere^{4,10,11}. P values refer to Mann–Whitney tests.

bystander fish should rise in the experimental situation but not in the controls. We found that 30 min after witnessing fights among conspecifics, the amounts of both testosterone and 11-ketotestosterone were significantly raised in the urine of the bystander, with 11-ketotestosterone remaining above basal concentrations for a further 3 hours after the fight (Fig. 1a, b). By contrast, in the control situation there was a continuous decline in both steroids, which is largely explained by the normal daily variation in basal urinary androgen, being highest in the early morning, at the start of phase 3, and then decreasing during the day (Fig. 1c).

Our results indicate that the endocrine system of spectators responds to social

interactions in which they do not themselves participate. Androgen levels pertaining in combatants 30 min after the start of a fight are a good predictor of the individual's position in the final hierarchy⁴, so the presence of increased concentrations in spectators suggests that these hormones mediate changes required for sharpening awareness and readiness to challenge. The potential adaptive benefits of increasing androgen levels during social challenge include a positive effect on cognitive tasks, promoting an animal's success socially by enhancing social attention, learning and memory processes^{5,6}.

Animals use information from previous interactive experiences to adjust their behaviour in subsequent social situations (as evidenced by the winner–loser effect⁷, eavesdropping⁸ and audience effects⁹, for example). Androgens are the likely mediators of these effects, acting by modulation of cognitive mechanisms underlying animal communication. We conclude that the 'challenge hypothesis' can be extended to bystander individuals. Androgens may eventually be viewed not only as sex steroids, but also as competition hormones that respond to the social environment and prepare the individual for competitive situations, rather as corticosteroids are considered as stress hormones in addition to their everyday function as metabolic regulators.

Rui F. Oliveira*, Marco Lopes*‡, Luis A. Carneiro*, Adelino V. M. Canário†

*Unidade de Investigação em Eco-Etologia, Instituto Superior de Psicologia Aplicada, Rua Jardim do Tabaco 44, 1100 Lisbon, Portugal
e-mail: ruiol@ispa.pt

†Centro de Ciências do Mar, Universidade do Algarve, Campus de Gambelas, 8000 Faro, Portugal

‡Present address: Primate Cognitive Neuroscience Laboratory, Center for the Behavioral Sciences, William James Hall, Harvard University, 33 Kirkland Street, Cambridge, Massachusetts 02138, USA

1. Pusey, A. E. & Packer, C. in *Behavioural Ecology* 4th edn (eds Krebs, J. R. & Davies, N. B.) 254–283 (Blackwell, Oxford, 1997).
2. Wingfield, J. C., Hegner, R. E., Duffy, A. M. & Ball, G. F. *Am. Nat.* **136**, 829–846 (1990).
3. Borg, B. *Comp. Biochem. Physiol. C* **109**, 219–245 (1994).
4. Oliveira, R. F., Almada, V. C. & Canário, A. V. M. *Horm. Behav.* **30**, 2–12 (1996).
5. Andrew, R. J. & Rogers, L. J. *Nature* **237**, 343–346 (1972).
6. Cynx, J. & Nottebohm, F. *Proc. Natl Acad. Sci. USA* **89**, 1376–1378 (1992).
7. Chase, I. D., Bartolomeo, C. & Dugatkin, L. *Anim. Behav.* **48**, 393–400 (1994).
8. Oliveira, R. F., McGregor, P. K. & Latruffe, C. *Proc. R. Soc. Lond. B* **265**, 1045–1049 (1998).
9. Gyger, M., Karakashian, S. J. & Marler, P. *Anim. Behav.* **34**, 1570–1572 (1986).
10. Scott, A. P. & Sumpter, J. P. *Gen. Comp. Endocrinol.* **73**, 46–58 (1988).
11. Scott, A. P., Mackenzie, D. S. & Stacey, N. E. *Gen. Comp. Endocrinol.* **56**, 349–359 (1984).

Semiconductor devices

Light-emitting diodes as chemical sensors

Semiconductor devices¹ make good chemical sensors, being responsive, selective and compact. Some gas analytes, for example, can be photometrically detected by the electroluminescence they absorb from semiconductor light-emitting diodes^{2,3}. Here we show that the intensity of this electroluminescence can be modulated in another way entirely, namely as a result of the adsorption of analyte onto the surface of the light-emitting diode, which affects its radiative efficiency by altering its surface-state structure and causes rapid, stable and reversible changes in electroluminescent intensity. Such devices are complementary to existing metal oxide-based resistance sensors, offering the advantages of low-power, room-temperature operation and ready integration into array-based devices.

Light-emitting diodes (LEDs) designed for high optical efficiency employ thin quantum-well active regions^{4,5}. However, the LED-based chemical sensors we describe here are designed to optimize the active layer's surface interactions so that adsorption effects on carrier recombination processes are accentuated. Such chemical adsorption can alter the nature and energy distribution of surface states that affect surface recombination velocity and radiative efficiency^{6,7}.

We used a double-heterostructure LED (Fig. 1a), grown by low-pressure metal-organic chemical vapour deposition. This has a conventional p-type/n-type junction arrangement, consisting of an active layer of undoped $\text{In}_{0.49}\text{Ga}_{0.51}\text{P}$, and p-type and n-type cladding layers of $(\text{Al}_{0.5}\text{Ga}_{0.5})_{0.51}\text{In}_{0.49}\text{P}$, grown to be lattice-matched to GaAs substrates; a highly doped p-type GaAs contact layer completes the structure. Each cladding layer is 0.5 μm thick and the active-layer thickness was varied from 50 to 500 nm. We processed the LED structures into broad-stripe ($500 \times 1,000 \mu\text{m}^2$) devices with cleaved (110) output facets.

The experimental set-up is adapted from semiconductor photoluminescent studies⁸. When forward-biased, the LEDs give red electroluminescence with band maxima at about 670 nm. Packaged devices were operated continuously at low and constant current levels (around 20 mA) to minimize heating and superluminescence. Electroluminescence experiments were carried out in the presence of five optically transparent analyte gases: ammonia (NH_3), methylamine (NH_2CH_3), dimethylamine ($\text{NH}(\text{CH}_3)_2$), trimethylamine ($\text{N}(\text{CH}_3)_3$) and sulphur dioxide (SO_2). A vacuum (10^{-3} torr) served as the reference ambient.

Figure 1b shows a typical trace of electroluminescence intensity against time for a

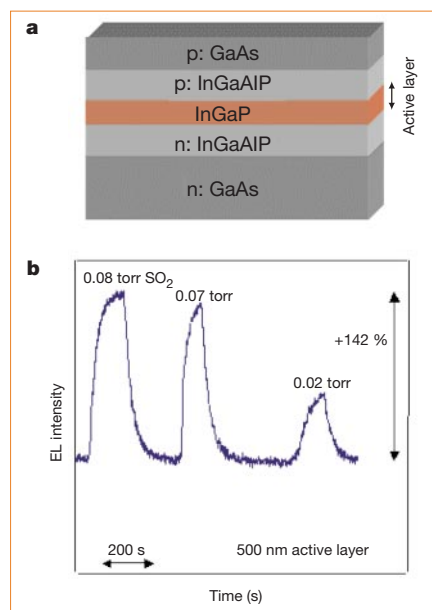


Figure 1 Diode structure and the response of its electroluminescence to sulphur dioxide. **a**, Double heterostructure p-n junction diode. Electrical contact is made to the top and bottom surfaces with metal films (layer dimensions not drawn to scale). **b**, Changes in electroluminescence intensity at 670 nm of a double-heterostructure diode upon exposure to SO_2 gas, relative to a vacuum ambient. The partial pressure of SO_2 is indicated; the sample is returned to the vacuum ambient between each exposure.

LED exposed to SO_2 . Enhancement increases with active-layer thickness, reaching several hundred per cent for the double heterostructures with the thickest active layer. The reversible responses (Fig. 1b) demonstrate the potential for online sensing with these structures. All of the analyte-diode structure combinations displayed adsorption and desorption times of less than a few minutes at room temperature. For all analytes, enhancements are seen at pressures as low as about 0.01 torr, with the response saturating at around 0.1–1 torr. The electroluminescent response is reproducible even after weeks of storage in air.

This adsorbate-induced modulation of electroluminescent intensity from simple double-heterostructure diodes illustrates their application as transducers that can couple semiconductor surface chemistry to an optical signal. The remarkable features of these structures (robust, inexpensive, small, low power consumption, compositionally tunable emission spectra) could allow arrays to be used to identify a variety of analytes. Highly sensitive, ultrasmall LEDs with large surface-to-volume ratios should then be easily integrated with diodes that serve as photodetectors, forming monolithic chemical sensors. From photoluminescent studies of semiconductors^{9,10} and from electroluminescence of molecular diodes¹¹, it should be possible to customize the selectivity, sensitivity and speed of the electroluminescent response by surface modification.

Albena Ivanisevic*, Jeng-Ya Yeh†, Luke

Mawst†, Thomas F. Kuech‡, Arthur B. Ellis*

Departments of *Chemistry, †Electrical and Computer Engineering and ‡Chemical Engineering, University of Wisconsin, Madison, Wisconsin 53706, USA

- Janata, J. & Josowicz, M. *Acc. Chem. Res.* **31**, 241–248 (1998).
- Krier, A., Sherstnev, V. V. & Gao, H. H. *J. Phys. D* **33**, 1656–1661 (2000).
- Ohyama, T. *et al. Sensors and Actuators B* **64**, 142–146 (2000).
- Ponce, F. A. & Bour, D. P. *Nature* **386**, 351–359 (1997).
- Stringfellow, G. B. in *High Brightness Light Emitting Diodes* Vol. 48 (eds Stringfellow, G. B. *et al.*) 1–45 (Academic, New York, 1997).
- Seker, F., Meeker, K., Kuech, T. F. & Ellis, A. B. *Chem. Rev.* **100**, 2505–2536 (2000).
- Boroditsky, M. *et al. J. Appl. Phys.* **87**, 3497–3504 (2000).
- Lorenz, J. K., Kuech, T. F. & Ellis, A. B. *Langmuir* **14**, 1680–1683 (1998).
- Kelly, M. T., Chun, J. K. M. & Bocarsly, A. B. *Nature* **382**, 214–215 (1996).
- Harper, J. & Sailor, M. J. *Anal. Chem.* **68**, 3713–3717 (1996).
- Kunugi, Y. *et al. Chem. Mater.* **10**, 1487–1489 (1998).

Public health

Screening slaughtered cattle for BSE

The systematic testing of slaughtered cattle aged over 30 months, or alternatively their elimination from the food chain, is an important component of a package of measures introduced in the European Union on 1 January 2001 to combat bovine spongiform encephalopathy (BSE) and protect human health. Here we explore the analytical limit of a rapid test designed to detect the abnormal prion protein associated with BSE in bovine brain and find that it is comparable to the limit of detection of infectivity in the conventional mouse bioassay¹, which is impractical for systematic screening. The sensitivity of the biochemical test allows it to be used as a viable alternative to the destruction of all carcasses of cattle slaughtered after 30 months of age. Additional work is required to compare this analytical sensitivity with the diagnostic sensitivity of the test under conditions of routine post-mortem BSE diagnosis and surveillance.

In May 1999, the European Commission evaluated four tests² based on the detection of PrP^{res} (also known as PrP^{Sc}), the abnormal protease-resistant form of the prion protein that is the only known molecular marker for BSE and related prion disorders. We have used a commercial version of the most analytically sensitive of these four tests (Bio-Rad version of the CEA test 'D') on serial dilutions of three samples taken from an undiluted homogenate of brainstem material pooled from several BSE cases and previously assayed for infectivity by titration in mice. The results in Fig. 1a show good agreement between the triplicate samples and a limit to detection of the absorbance signal in the assay at a 1/1,000 dilution of the infected brain homogenate.

In the mouse bioassay, the infectivity

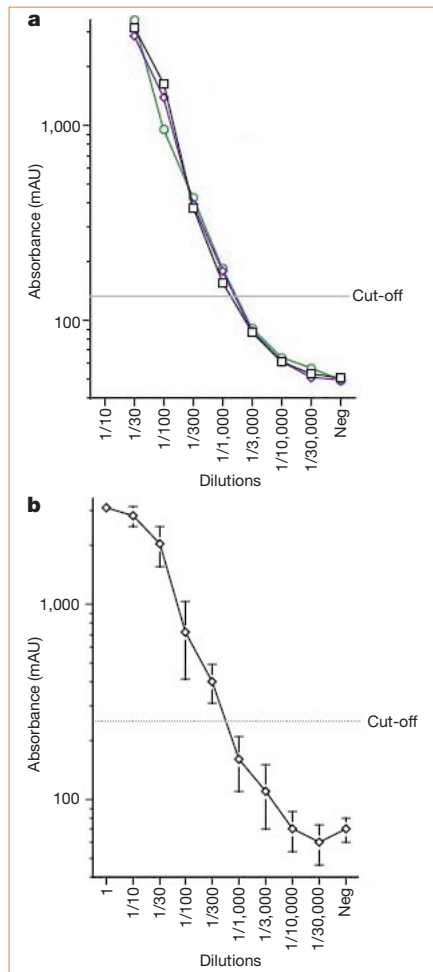


Figure 1 Testing *in vitro* for protease-resistant prion protein (PrP^{res}) in BSE-infected brain tissue. **a**, Triplicate samples from pooled infected brainstem homogenate were assayed using a commercial BSE-testing kit (Bio-Rad) based on sandwich immunoassay of rapidly purified PrP^{res} by enzyme-linked immunosorbent assay (ELISA)^{2,3}. Samples were rehomogenized at 20% w/v in the buffer of the Bio-Rad purification kit. Serial dilutions were made using a pool of 20% w/v BSE-negative brain homogenate prepared in the same conditions. Testing was performed according to the manufacturer's indications. The equivalent of 33 mg of brain was loaded in each ELISA well. We used the absorbance limit recommended by the manufacturer (defined as the background absorbance of the ELISA buffer R3 + 0.090 absorbance units (AU) = 0.133 AU) to determine the cut-off sensitivity of the test. All dilutions up to 1/1,000 were above the cut-off for all three samples. The signal detected at the 1/3,000 dilution was below the recommended cut-off. The titre of the inoculum was 10^{3.3} LD₅₀ g⁻¹, or 66 LD₅₀ units per 33 mg of brain (equivalent to the amount loaded in one ELISA well); the dilution 1/1,000 corresponds to 0.066 LD₅₀ units per ELISA well. 'Neg' indicates normal bovine brain material. **b**, PrP detection in diluted homogenates tested under blind conditions during the EC evaluation of May 1999 (20 samples per dilution; errors bars correspond to standard deviation)¹⁻³. The commercial kit used in **a** employs the enzyme peroxidase, which gives a lower R3 background absorbance than the acetylcholinesterase used in the earlier tests. The sensitivity cut-off here is therefore higher (0.250 AU), corresponding to 2.5 times the mean of the negative samples analysed (2.5 × 100 mAU = 250 mAU). This cut-off was consistent with a 100% specificity on a series of 1,064 negative bovine samples from New Zealand; 18/20 homogenates were detected as positive at a dilution of 1/300, and 1/20 at a dilution of 1/1,000 with this homogenate¹ (infectivity 10^{3.1} LD₅₀ g⁻¹, or 35 LD₅₀ per 28 mg of brain (equivalent to the amount loaded in one ELISA well); the 1/300 dilution corresponds to the equivalent of 0.1 LD₅₀ per ELISA well).

titre of the reference material inoculated into mice was 1 LD₅₀ unit per 0.5 mg infected bovine brain (where the dose corresponding to one LD₅₀ unit has a 50% chance of killing the mouse); the limit of detection of the biochemical test corresponds to 0.033 mg of this BSE-positive bovine brain material (Fig. 1, legend). Direct comparison with results from titration in mice (Table 1) shows that the highest dilution of material for which infection can be detected in the

bioassay also gives a positive result in the biochemical test (1/1,000; 3 of 3 samples positive). The limit of detection of the test is therefore well below 1 LD₅₀ unit in mice. This is a critical finding because most information on infectivity refers to data from this conventional mouse model, although the sensitivity of detection of infectivity should be improved as transgenic mouse models become available.

The calibration curves shown in Fig. 1a

are consistent with results obtained using different diluted homogenates of similar infectivity and tested blind during early evaluation by the European Commission²⁻⁴ (Fig. 1b). This provides an indication of the reproducibility of this PrP^{res} detection test and suggests a direct relationship between infectivity titre and PrP^{res} concentration.

The relevance of making serial dilutions to evaluate the sensitivity has been confirmed on preclinical samples containing various concentrations of BSE agent using central nervous system tissue obtained in a pathogenesis study conducted in the United Kingdom⁵ (our unpublished results).

Our results show that this PrP^{res}-detection test on brain material offers a level of assurance regarding the detection of the presence of infection that is at least comparable with that provided by the conventional mouse bioassay. Although bovines are regarded as the most sensitive recipient species for BSE, the oral route in cattle is less efficient in establishing infection with the BSE agent than is direct injection into mouse brain: the differential factor is currently estimated¹ at around 100 (on the grounds that 1 g brain titrating 10^{3.3} LD₅₀ g⁻¹ can theoretically kill, with a probability of 50%, 2,000 mice injected intracerebrally or 20 bovines contaminated orally). It is reasonable to presume that humans contaminated by the oral route are partially protected by the species barrier and so are less sensitive than cattle to the BSE agent (although this differential factor is not known). Consequently, risk assessment generally assumes that humans contaminated by the oral route are less sensitive to the BSE agent than mice inoculated intracerebrally.

Routine application of the test should help eliminate previously undiagnosed BSE-affected animals from the food chain and provide tighter epidemiological monitoring of the BSE epidemic.

J. P. Deslys*, E. Comoy*†, S. Hawkins‡, S. Simon§, H. Schimmel||, G. Wells‡, J. Grassi§, J. Moynagh¶

*CEA, Service de Neurovirologie, DRM/DSV, BP6 92265 Fontenay-aux-Roses, France
e-mail: jpdesslys@cea.fr

†Bio-Rad Life Sciences Laboratories, Hercules, California 94547, USA, and Marnes-la-Coquette 92430, France

‡Veterinary Laboratories Agency, New Haw, Addlestone, Surrey KT15 3NB, UK

§CEA, Service de Pharmacologie et d'Immunologie, DRM/DSV, Saclay, France

||Joint Research Centre, Institute for Reference Materials and Measurements, Retieseweg, B2440 Geel, Belgium

¶Directorate General Health and Consumer Protection, European Commission, Rue Belliard 232, 1049 Bruxelles, Belgium

1. European Commission in Oral Exposure of Humans to the BSE Agent: Infective Dose and Species Barrier 31 (Scientific Steering Committee meeting, 13–14 April 2000).

2. Moynagh, J. & Schimmel, H. *Nature* **400**, 105 (1999).

Table 1 ELISA detection of PrP^{res}

Dilution	Sample 1	PrP ^{res} detection Sample 2	Sample 3	Infectivity (dead/total)
1/10				17/18
1/30	3,166	2,880	3,466	
1/100	1,618	1,318	958	15/17
1/300	378	396	429	
1/1,000	155	179	186	1/14
1/3,000	87	88	91	
1/10,000	61	61	64	0/7
1/30,000	53	51	57	
1/100,000	-	-	-	0/9
Neg	51	49	49	NT
ELISA buffer	43	43	43	
Cut-off	133	133	133	

Absorbance values for each sample correspond to those shown in Fig. 1a. To determine infectivity, reference BSE-infected brain homogenate was titrated in Rll mice. Serial dilutions of brain were prepared and each of 20 mice per group was inoculated with a single dilution (20 µl by intracerebral route + 100 µl by intraperitoneal route per animal). Results are expressed as the number of mice developing the disease/number of mice surviving at the time when the first mouse developed the disease. (The denominator of the incidence of mice succumbing to disease is corrected to eliminate mice lost to the assay before the minimum incubation period started at which they could have developed disease; in groups where no infectivity was detected, the denominator is expressed as the number of mice surviving to the completion of the assay at 700 days.) NT, not tested. Neg, normal bovine brain.

3. European Commission *The Evaluation of Tests for the Diagnosis of Transmissible Spongiform Encephalopathy in Bovines* (http://europa.eu.int/comm/food/fs/bse/bse12_en.html).
4. Grassi, J. et al. *Arch. Virol.* **16** (suppl.), 197–205 (2000).
5. Wells, G. A. H. et al. *Vet. Rec.* **142**, 103–106 (1998).

Genome evolution

Gene capture in archaeal chromosomes

Free genetic elements can be readily integrated into bacterial chromosomes, but so far, with the exception of one virus, there has been no evidence that this happens in Archaea — the other domain of microorganisms. Here we show that site-specific integration of different genetic elements into archaeal chromosomes is a general phenomenon, albeit rare, which requires an archaeal integrase and produces a partitioned integrase gene in the chromosome. The process is distinct from bacterial mechanisms and has implications for how horizontal gene transfer might occur across the boundaries of the domains of life.

Many bacterial integrases belong to a superfamily of tyrosine recombinase enzymes¹ and mediate site-specific recombination between small extra-chromosomal DNA elements and chromosomes², facilitating horizontal gene transfer between bacteria². The only known archaeal integrase effects site-specific and reversible integration of the SSV1 virus into the chromosome of *Sulfolobus shibatae*^{3,4} (Fig. 1a, top).

We have previously identified a region of the *Sulfolobus solfataricus* genome that shows high sequence similarity to the *Sulfolobus* plasmid pHEN7 and to other members of the pRN plasmid family^{5–7}. This chromosomal region is flanked by open-reading

frames (ORFs) with extensive sequence similarity to the amino-terminal (66 amino acids) and carboxy-terminal (269 amino acids) regions of the SSV1 integrase (Fig. 1a). The former, Int(N), overlaps with the downstream half of a gene encoding tRNA^{Val} and shows 71% sequence identity to the 45-base-pair target site (*att*) of the SSV1 integrase⁴ (within the downstream half of a tRNA^{Arg} gene), whereas the latter, Int(C), contains an identical *att* site (Fig. 1a, b).

Excision from the chromosome by recombination at the direct *att* repeat and circularization creates a plasmid, pXQ1, containing an intact integrase gene and one copy of the *att* site, similar to that of the SSV1 virus⁴. Thus, with the aid of the encoded integrase, pXQ1 could have integrated into the chromosome. As neither the free form of pXQ1 nor the empty chromosomal site of *Sulfolobus* were detectable by polymerase chain reaction (our unpublished results), we infer that the absence of an intact integrase gene precludes integrase-mediated excision from the chromosome.

A search for more homologues of the SSV1 integrase encoded in the chromosome of *S. solfataricus*⁵ revealed three Int(N)s and one Int(C). The sequences encoding Int(N) overlap with downstream halves of tRNA genes and all the sequences contain *att* sites. By analogy to the SSV1 virus, a genetic element XQ2 (67.7 kilobases) can be formed by recombination and excision at the *attL* and *attR* sites (Fig. 1a, b).

Four of the ORFs encoded in XQ2 were assigned to genes encoding the enzymes dTDP-glucose-4,6-dehydratase, glucose-1-phosphate-thymidyl transferase, dTDP-4-dehydrorhamnose reductase and dTDP-4-dehydrorhamnose-3,5-epimerase, all of which are involved in central metabolism,

and three of which are generally clustered in one operon. However, there are two additional copies of dTDP-glucose-4,6-dehydratase and another copy of each of the other enzymes encoded within the *S. solfataricus* chromosome⁵ that show 76–84% sequence identity and 86–96% similarity. It is therefore likely that XQ2 entered the chromosome, aided by its own integrase, to produce a gene-capture event.

We searched the complete genome sequences of seven other archaea for integrase homologues⁸. Those from the euryarchaeote *Pyrococcus horikoshii*⁹ and the crenarchaeote *Aeropyrum pernix*¹⁰ each had two Int(N), overlapping downstream halves of tRNA genes, and two and one copy of Int(C), respectively. The genome of *Pyrococcus* OT3 has two partitioned integrase genes¹¹. Complementary fragments of the integrase genes contain *att* sites and border regions of 21.5 and 4 kilobases in *P. horikoshii* and 17.5 kilobases in *A. pernix*. Compatible with the idea that these are inserts from unknown organisms, none of their ORFs reveals any hits in the GenBank/EMBL databases⁸. We have detected a total of seven partitioned integrase genes, on average one per genome. Their sequences vary markedly in the amino-terminal region but are less variable in the carboxy-terminal region (22–33%/41–51% identity/similarity), indicating that other integrase genes may still remain to be discovered.

We conclude that chromosomal integration in archaea differs from that in bacteria in producing partitioned integrase genes and that integration can be reversed only if the intact integrase is produced⁴. ‘Curing’ the cell of the free genetic element carrying the intact integrase gene can thus lead to gene capture by the archaeal chromosome. Although archaeal integrases differ from bacterial integrases in lacking a conserved motif¹ and in having *att* sites in the coding region, they may still facilitate gene transfer between archaea, bacteria and eukaryotes¹².

Qunxin She*, Xu Peng*, Wolfram Zillig†, Roger A. Garrett*

*Microbial Genome Group, Institute of Molecular Biology, University of Copenhagen, Sølvgade 83H, DK-1307 Copenhagen K, Denmark
†Max-Planck-Institut für Biochemie, D-82152 Martinsried, Germany

1. Nunes-Düby, S. E. et al. *Nucleic Acids Res.* **26**, 391–406 (1998).
2. Rowe-Magnus, D. A. & Mazel, D. *Curr. Opin. Microbiol.* **2**, 483–488 (1999).
3. Reiter, W.-D. & Palm, P. *Mol. Gen. Genet.* **221**, 65–71 (1990).
4. Muskhelishvili, G., Palm, P. & Zillig, W. *Mol. Gen. Genet.* **237**, 334–342 (1993).
5. She, Q. et al. *DNA Sequence* **11**, 183–192 (2000).
6. Kletzin, A., Lieke, A., Ulrich, T., Charlebois, R. L. & Sensen, C. W. *Genetics* **152**, 1307–1314 (1999).
7. Peng, X., Holz, I., Zillig, W., Garrett, R. A. & She, Q. *J. Mol. Biol.* **303**, 449–454 (2000).
8. Altschul, S. F. et al. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
9. Kowarabayasi, Y. et al. *DNA Res.* **5**, 55–76 (1998).
10. Kowarabayasi, Y. et al. *DNA Res.* **6**, 83–101 (1999).
11. Makino, S., Amano, N., Koike, H. & Suzuki, M. *Proc. Jpn. Acad. Ser. B* **75**, 166–171 (1999).
12. Groth, A. C., Olivares, E. C., Thyagarajan, B. & Calos, M. P. *Proc. Natl Acad. Sci. USA* **97**, 5995–6000 (2000).

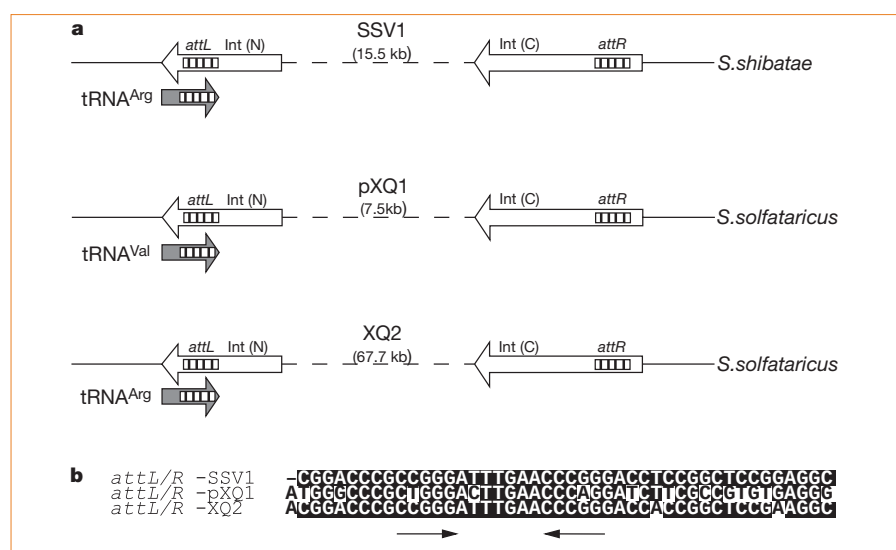


Figure 1 Site-specific integration of genetic elements into archaeal chromosomes. **a**, Organization of the SSV1 virus and the putative genetic elements pXQ1 and XQ2 in the chromosomes of *S. shibatae* and *S. solfataricus*. The relative positions of the partitioned integrase genes, the *att* sites (45 base pairs) and the overlapping tRNA genes are indicated. **b**, Alignment of the direct repeats *attL* and *attR* from virus SSV1 and from pXQ1 and XQ2. Nucleotides are shaded if two or more positions are identical. Arrows indicate a 5-base-pair inverted repeat.

The architecture of active zone material at the frog's neuromuscular junction

Mark L. Harlow, David Ress, Arne Stoschek, Robert M. Marshall & Uel J. McMahan

Department of Neurobiology, Stanford University School of Medicine, Stanford, California 94305, USA

Active zone material at the nervous system's synapses is situated next to synaptic vesicles that are docked at the presynaptic plasma membrane, and calcium channels that are anchored in the membrane. Here we use electron microscope tomography to show the arrangement and associations of structural components of this compact organelle at a model synapse, the frog's neuromuscular junction. Our findings indicate that the active zone material helps to dock the vesicles and anchor the channels, and that its architecture provides both a particular spatial relationship and a structural linkage between them. The structural linkage may include proteins that mediate the calcium-triggered exocytosis of neurotransmitter by the synaptic vesicles during synaptic transmission.

At synapses throughout the nervous system, one or more areas on the cytoplasmic surface of the plasma membrane of presynaptic nerve cells are studded with distinctive aggregates of proteins. The aggregates can only be observed by electron microscopy (EM), and they are viewed best in sections from fixed, plastic-embedded tissue, where they stain with heavy metals¹. They are confined to the portion of the presynaptic membrane that faces the narrow cleft between the pre- and postsynaptic cell, and some of the presynaptic cell's numerous synaptic vesicles, which contain neurotransmitter molecules, are docked on (that is, held at) the membrane next to them. When an electrical impulse in the presynaptic nerve cell arrives at the synapse, it causes calcium channels in the presynaptic membrane near to the protein aggregates to open, thereby permitting calcium ions to enter the cell. The influx of calcium triggers events that cause the membrane of the docked vesicles to fuse with the presynaptic membrane, and the subsequent exocytosis of the neurotransmitter into the synaptic cleft^{2–4}. Thus, the protein aggregates are situated at the active zones of the presynaptic membrane, the sites where initial events in the rapid neurotransmitter-mediated signalling between the pre- and postsynaptic cells take place. Since their detection⁵, the protein aggregates have been referred to in a variety of ways including membrane thickenings and presynaptic dense projections as well as active zone material (AZM), the term we use here.

The proximity of AZM to docked synaptic vesicles and calcium channels suggests that it is physically connected to one or both structures, and raises questions about the possible function of such connections^{6–8}. We used electron microscope tomography (EMT) to examine the structure and associations of AZM in tissue sections from neuromuscular junctions. Transmission EM provides two-dimensional projections of specimens, which are captured on film or CCD (charge-coupled device). Conventionally, data analysis is done on these projections, but the level of detail is restricted by the absence of depth information. In particular, components of the AZM are indistinct and their relationships are uncertain even in the thinnest sections that can be routinely cut (30–50 nm). Electron microscope tomography, which has been developed mostly over the last decade, provides three-dimensional reconstructions of specimens by using a series of two-dimensional projections made at different tilt angles⁹. The increased spatial resolution obtained from three-dimensional reconstructions over two-dimensional projections—more than an order of magnitude along the z-axis—has already been useful for markedly enhancing understanding of certain macromolecules, viruses and cellular organelles in *in vitro* preparations^{10–14}. It has also been used for determining in stained tissue sections the shapes and relationships of subcellular structures

that have a high signal/noise ratio and a simple geometry^{15–18}. Because the components of AZM have a complex configuration and are of moderate to low contrast in tissue sections, we found it necessary to combine advanced image restoration and new three-dimensional segmentation methods for our analysis. We provide a detailed account of the organization and connections of AZM, and a refined hypothesis as to its function.

Overviews

Conventional two-dimensional EM has shown that each patch of AZM at the frog's neuromuscular junction is elongate, with its long axis lying parallel to the presynaptic membrane^{3,19} (Fig. 1a, b). It is frequently 1–2 μm long and about 75 nm wide, extending 50–75 nm from the presynaptic membrane into the terminal's cytoplasm. On each flank there is a row of docked vesicles, and where the AZM is adjacent to the presynaptic membrane, the membrane curves outward forming a 'ridge'¹⁹. On each slope of the ridge, as seen by EM on replicas of freeze-fractured specimens, the membrane contains a linear array of macromolecules that parallels the long axis of the AZM. The array may include multiple species of proteins, but evidence indicates that the proteins that form calcium channels are concentrated within it^{20–22}. Thus, at this synapse, the calcium channels are apparently concentrated in two specific portions of the presynaptic membrane directly beneath the AZM, each portion parallel and close to one of the rows of docked vesicles flanking the material (Fig. 1a, b).

Figure 1c is a two-dimensional EM image of an active zone in a 50-nm-thick tissue section from a frog's neuromuscular junction. It is similar to the images used for the scheme in Fig. 1a, b. It is also the 0° image of a data set of two-dimensional tilt images, called MPI-10, that we used to reconstruct the section's volume. The plane of the section is nearly transverse to the AZM. The AZM is flanked by docked vesicles, and its superficial surface follows the contour of the presynaptic membrane where it forms the active zone ridge. Components of the AZM are evident, but their size, shape and relationships are indeterminate because of the absence of depth information.

There is no uncertainty, however, about the components of AZM in two-dimensional images of 0.5-nm (2 voxel) slices through the reconstructed volumes from MPI-10 and from another active zone, MPI-9 (Fig. 1d, e). Individual components are spatially resolved in three dimensions, so that their size, shape and associations can be determined accurately. Serial slices through the entire volume of the active zones revealed that the components were generally elongate. One end of some components was continuous with the membrane of the docked vesicles. In such cases, the otherwise nearly smooth

curvature of the vesicle membrane was distorted toward the components (Fig. 1d, e). These distortions would be expected if the components were connected to the vesicles, and the connections had withstood some tension generated either *in vivo* or during preparation for microscopy.

A more coherent perspective of the AZM, synaptic vesicles and presynaptic membrane in MPI-10 and MPI-9 was obtained by segmenting and surface-rendering these structures. The three-dimensional views showed that much of the four docked vesicles had been included in the tissue sections (Fig. 1f–i). They also showed that the AZM was widely distributed over the slopes of the active zone ridge, that it formed many connections with each vesicle, and that the components were, in general, interconnected.

Connections to docked synaptic vesicles

Examination of the serial volume slices from MPI-9 and MPI-10 show that the number of components connected to the four partial vesicles ranged from seven to twelve, with each component forming only one connection. In two cases, two components seemed to be conjoined at their sites of connection to the vesicles, but in all other cases the components contacted the vesicles separately. We used surface renderings to examine the distribution of connection sites between the vesicles and AZM. To expose the connections, we made partial segmentations that included only the short portions (~1 nm) of the components where they contacted the vesicles (Fig. 2a–d). Connection sites were widely distributed on the half of the vesicles facing the AZM. Several connections were 10–15 nm

from the vesicle–presynaptic membrane interface, where vesicle fusion with the presynaptic membrane takes place during synaptic transmission. Cytoplasmic structures other than those comprising the AZM also were in contact with docked vesicles in MPI-9 and MPI-10 (Fig. 1f–i). However, there were only three or four such contacts per vesicle, and they were primarily on the half of the vesicles that faced away from the AZM.

Beams and ribs

To determine whether any of the components that contacted the docked synaptic vesicles near the vesicle–presynaptic membrane interface were also associated with the macromolecule-containing regions of the presynaptic membrane along the slopes of the active zone ridge, we undertook a detailed examination of the architecture of the AZM within 15 nm of the membrane.

All of the components in the first 15 nm of AZM in MPI-10 and MPI-9 were elongate (Figs 3a–d and 4a). We initially identified two types on the basis of their position and orientation, which we call beams and ribs. The beams lay near the midline of the presynaptic ridge, nearly parallel to the ridge's long axis. Ribs extended along the slopes of the ridge, nearly orthogonal to its long axis. In MPI-10, six of the seven ribs contacted the synaptic vesicles near the vesicle–presynaptic membrane interface (Figs 3c and 4a), and they accounted for all of the contacts made by AZM at the interface (Fig. 2a, b). The one rib that did not contact a vesicle was at the surface of the section, and may have contacted a vesicle that was not included in the section. Nearly all of the ribs from both slopes of the

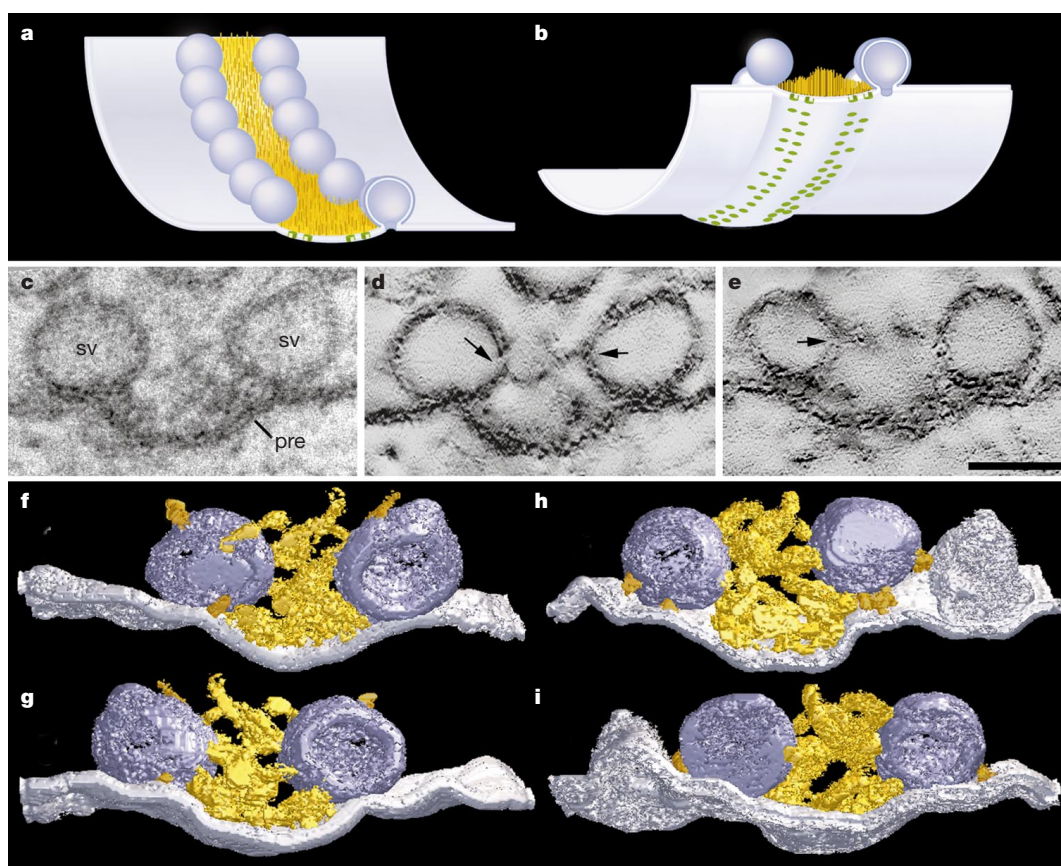


Figure 1 Two-dimensional and three-dimensional (2D and 3D, respectively) overviews. **a, b**, Schematized relationships of AZM based on conventional 2D EM. The colour code for AZM (gold), docked synaptic vesicles (dark blue), presynaptic membrane with active zone ridge (pale blue) and macromolecules along the slopes of the ridge (green) are the same in all figures, except where indicated. **c**, The 0° image of MPI-10 from a tissue section. AZM is indicated by an arrow. sv, docked synaptic vesicles; pre, presynaptic membrane. **d, e**, 0.5-nm volume slices from the reconstruction of MPI-10 (**d**) and MPI-9 (**e**). The smooth

curvature of the vesicles is distorted at the sites of contact (arrows) between AZM and vesicles, indicative of tension. Scale bar, 50 nm (**c–e**). **f–i**, 3D views of segmented, surface-rendered AZM, docked vesicles and presynaptic membrane in MPI-10 (**f**, front; **g**, back) and MPI-9 (**h**, front; **i**, back). The hummock in the presynaptic membrane in **h** and **i** is a vesicle of unknown type that is undergoing either exo- or endocytosis. Cytoplasmic structures associated with the synaptic vesicles but not part of the AZM are bronze.

ridge in MPI-10 also terminated on a single beam (Fig. 3c). In the two exceptions, ribs from opposite slopes terminated on a short structure at the surface of the section, which was probably part of a beam that was not fully included in the section. In MPI-9, seven of the eight ribs contacted a beam (Fig. 3d): the exception was interrupted near the beam. The plane of the tissue section for MPI-9 was intermediate to the transverse and horizontal planes of the AZM, so that the lateral portions of some of the ribs were not included in the section. However, each of the four ribs that could be traced to their lateral ends contacted a vesicle near the vesicle–presynaptic membrane interface (Fig. 3d), and these contacts accounted for all of the contacts made by AZM at the interface (Fig. 2c, d). Thus, for both MPI-9 and MPI-10, all of the components of AZM that contacted vesicles near the vesicle–presynaptic membrane interface were ribs, all ribs lay in close apposition to the presynaptic membrane along the slopes of the active zone ridge, and nearly all of the ribs were connected to a beam.

We observed many more ribs and beams in the first 15 nm of AZM in another tissue section that we reconstructed from data set UC-1. The tissue section was cut nearly parallel to the horizontal plane of the AZM, so that more of the active zone was included than in the sections used for MPI-9 and MPI-10, and data were collected at a lower magnification. Thirteen docked vesicles were associated with forty ribs that were nearly parallel to one another and orthogonal to the long axis of the active zone ridge (Fig. 3e, f). Seven beams lay parallel to the long axis of the ridge near its midline. Five beams were in series and two were paired, with the ends of adjacent ones overlapping and contacting one another (Fig. 3e–g). One end of all the ribs contacted a vesicle (Fig. 3f), and the other end of most ribs contacted a beam (Fig. 3e, f). For the exceptions, the gaps between the ribs and beams were of various lengths. Thus, the arrangement and the connectivity of ribs and beams in the relatively long stretch of AZM provided by UC-1 were the same as in MPI-9 and MPI-10.

We found the following characteristics of ribs and beams: those ribs terminating on each of the paired beams in UC-1 were shorter

than those terminating on the serial beams (Fig. 3e, f). This shows that ribs can have different lengths. In most cases, the ribs tapered as they approached their contact points with beams (Fig. 3). They also tapered as they approached their contact points with synaptic vesicles, but more acutely (data not shown). Indeed, the general thickness of the ribs was usually greatest where they were adjacent to the region of the presynaptic membrane that contained the linear array of macromolecules. Of the ten beams we observed in MPI-9, MPI-10 and UC-1, we could be certain of only two having their full length included in the samples (Fig. 3g). Both were ~75 nm long; although some of the uncertain beams were as long as these two ~75-nm beams, none were longer.

To learn whether there was a correlation between the frequency of the ribs along the active zone ridge and the frequency of macromolecules in the adjacent presynaptic membrane, we compared our images of rib–beam assemblies with images of the macromolecules in replicas from freeze-fractured neuromuscular junctions of the frog. The linear array of macromolecules on each slope of an active zone ridge in a previously published image⁴ is shown in Fig. 3h. Each array is an often-interrupted string of conspicuously large particles that have a nearly regular spacing characterized by a longitudinal frequency. The particles occur singly or in pairs, the latter characterized by a transverse frequency. The same image of the replica, overlaid with images of surface-rendered rib–beam assemblies from UC-1, shows a strong similarity between the longitudinal frequency of the particles and the frequency of ribs (Fig. 3i). We tested this correlation by determining the mean centre-to-centre spacing of continuous runs of particles (1,100 measurements; only one member of paired particles was included) in replicas²³ from five freeze-fractured neuromuscular junctions and comparing it with the mean midline-to-midline spacing of the ribs in the rib–beam assemblies from MPI-9, MPI-10 and UC-1 (45 measurements on surface renderings at points where the ribs overlay the slopes of the active zone ridge). The muscles were fixed with glutaraldehyde, and were taken from frogs of the same age and size as those used for tomography (see Methods). The spacing of the particles/macromolecules (17.2 ± 3.6 nm) was in good agreement with the spacing of the ribs (16.1 ± 3.4 nm).

Pegs

Examination of the serial volume slices from the reconstructions of MPI-9 and MPI-10 showed that the ribs were separated from the presynaptic membrane by a narrow gap (≤ 7 nm). At intervals, components of AZM extended across the gap, connecting the ribs to the regions of the membrane containing the macromolecules (Fig. 4a). We have called these components pegs. Segmentations of rib–beam assemblies in MPI-9 and MPI-10, together with a portion (~1 nm) of each peg where it connected to a rib, showed that each rib had one or two such connections (Fig. 4b, c). Segmentation of the presynaptic membrane in MPI-9 and MPI-10, together with a portion of each peg where it connected to the membrane, showed that the longitudinal and transverse frequencies of the peg–membrane connections along the active zone ridge were markedly similar to the longitudinal and transverse frequencies of the macromolecules in freeze-fracture replicas (Fig. 4d–f). Cytoplasmic proteins generally connect to membranes by binding to the membranes' proteins. Because the only discernible membrane proteins on the slopes of the active zone ridge that had a distribution similar to that of the pegs were the macromolecules, our findings indicate that all ribs are connected by means of pegs to most, if not all, of the macromolecules in the presynaptic membrane, and that each rib is so connected to at least one or two macromolecules.

Refined and extended hypothesis

Our study shows that the AZM at the frog's neuromuscular junction consists of elongate, interconnected components, and that some of these components are also connected to synaptic vesicles docked at

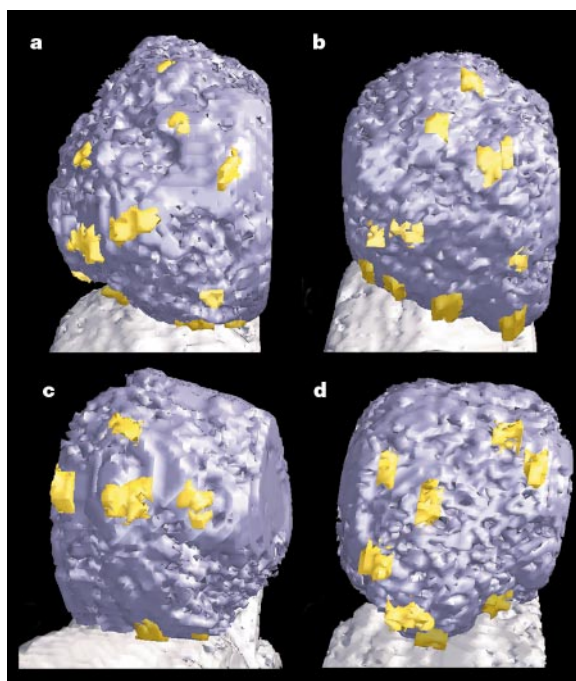


Figure 2 Sites of connections between components of AZM and docked vesicles. **a–d**, Median view of docked vesicles in MPI-10 (**a**, **b**) and MPI-9 (**c**, **d**) together with short portions of components (light and dark gold) marking the connections. The portions nearest to the presynaptic membrane (dark gold) are the ends of ribs.

the presynaptic membrane. The fact that each docked vesicle included in our samples was connected to many components provides strong evidence that the AZM, as previously suggested⁶, is involved directly in vesicle docking. We also show that some components of the AZM are connected selectively to the narrow regions of presynaptic membrane that contain a linear array of macromolecules, which includes calcium channels, and we provide evidence that most, if not all, of the macromolecules have such connections. Proteins clustered in the postsynaptic membrane of muscle fibres and in the plasma membrane of other cell types are known to be anchored to cytoplasmic proteins, which inhibit their lateral mobility²⁴. Accordingly, it is probable that components of the AZM that are linked to the macromolecules in the presynaptic membrane contain proteins that help to anchor calcium channels and any other membrane proteins that may also be in the linear

arrays^{25,26}.

Our detailed examination of AZM that lay within 15 nm of the presynaptic membrane revealed that the components have a highly ordered distribution, with there being three topologically distinct types: ribs, beams and pegs. In most cases, these three components were interconnected in a systematic way, and this assembly had specific connections to docked synaptic vesicles and regions of the presynaptic membrane that contained the macromolecules. In the few cases where there were gaps in and between components that were not apparent elsewhere, it is probable that such gaps were the result of our inability to adequately fix or stain the components during preparation for microscopy. Thus, we propose (Fig. 5) that for the first 15 nm of AZM at resting neuromuscular junctions, the typical architecture consists of beams that are linked to each other and to ribs, ribs linked to

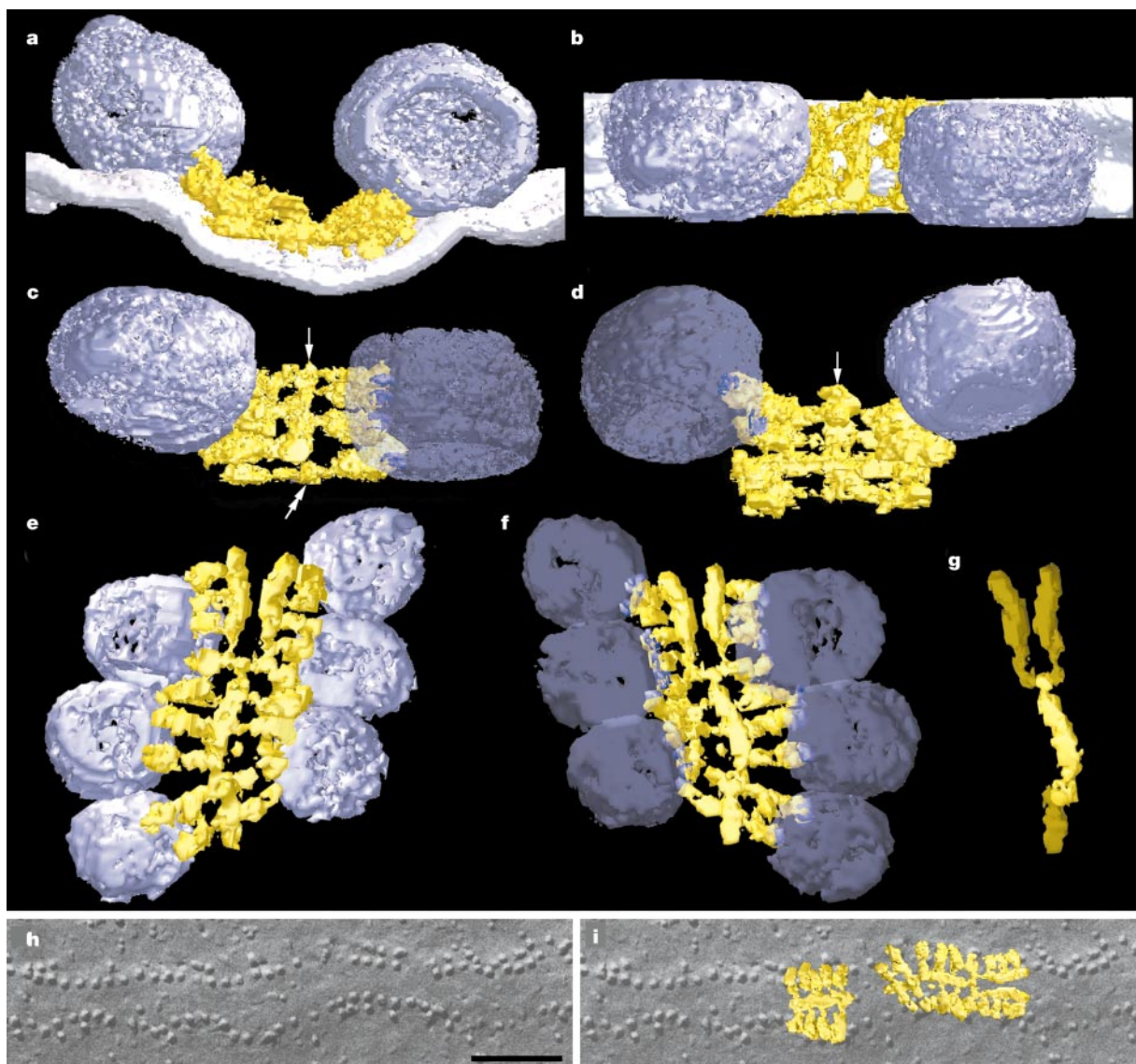


Figure 3 Distribution and associations of ribs and beams. **a, b**, Transverse and horizontal views, respectively, of the first 15 nm of AZM internal to the presynaptic membrane in MPI-10, together with the membrane and docked vesicles. **c**, View similar to that in **b** with the presynaptic membrane removed. We have made one vesicle transparent to show four complete ribs. Contact sites of these ribs with the transparent vesicle are indicated as bright blue patches. Most of the ribs contact the same beam (arrow), and two contact another beam (double arrow). **d**, View of ribs and a beam (arrow) in MPI-9, similar to that shown for MPI-10 in **c**. **e–g**, Horizontal views from the presynaptic membrane (**e**), and from the cytoplasm (**f, g**) of the first 15 nm of AZM and docked vesicles from a portion of

the reconstructed volume in UC-1. Numerous ribs extend between vesicles and beams (**e**). We have made the vesicles transparent to show points of rib-vesicle contact (bright blue) (**f**); only the beams are shown in **g**. One beam (light gold) can be traced to connections to other ribs at both of its ends. **h**, Replica⁴ from a freeze-fractured frog's neuromuscular junction showing a series of particles/macromolecules on each slope of a ridge. Scale bar, 100 nm. **i**, Same replica with surface-rendered rib-beam assemblies from UC-1 superimposed at the same scale as the replica, which shows the similarity in longitudinal frequency between ribs and macromolecules. (See Supplementary Information for movies.)

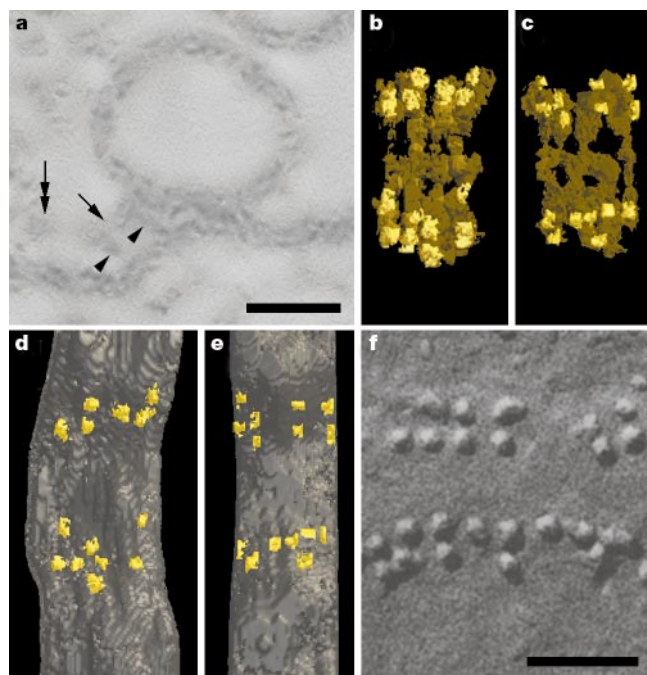


Figure 4 Arrangement and relationships of pegs. **a**, 2D projection of a 2.5-nm volume slice from MPI-10, transverse to the AZM, showing a rib (arrow) extending from a synaptic vesicle toward a beam (double arrow). Two pegs (arrowheads) extend from the rib to the presynaptic membrane. Scale bar, 25 nm. **b, c**, Horizontal views of surface-rendered rib-beam assemblies (dark gold) in MPI-10 and MPI-9, respectively, that also include short portions of pegs (light gold) marking the sites where the pegs contact the ribs. **d, e**, Horizontal views of surface-rendered presynaptic membranes in MPI-10 and MPI-9, respectively, that also include short portions of pegs (gold) marking the sites where the pegs contact the membranes. **f**, Portion of the replica of freeze-fractured presynaptic membrane shown in Fig. 3h that is scaled for comparing the longitudinal and transverse frequencies of macromolecules in the membrane to the longitudinal and transverse frequencies of peg-membrane contacts in **d** and **e**. Scale bar, 50 nm.

docked vesicles and pegs, and pegs linked to the macromolecules in the presynaptic membrane. This leads to the broader hypothesis that the assembly of ribs, beams and pegs, augmented by the remainder of the AZM, acts as a scaffold that helps to maintain the linear arrangement of, and spacing between, docked vesicles and the macromolecules.

At least some of the pegs and ribs may be involved directly in regulating the fusion of docked vesicles with the presynaptic membrane during synaptic transmission. On the basis of the amount of the calcium channel protein that is cytoplasmic²², each channel may make up part or all of a peg, and even extend into part of a rib. Ribs may also contain proteins that mediate the effects of calcium on vesicle fusion. Evidence indicates that the cytoplasmic domains of certain proteins in the membrane of docked vesicles are bound to the cytoplasmic domains of certain other proteins in the presynaptic membrane, and the association of these so-called SNARE proteins brings the vesicles and presynaptic membrane into direct apposition^{27,28}, perhaps into partial fusion²⁹. Moreover, the cytoplasmic domains of one of the presynaptic membrane's SNARE proteins, syntaxin, and another vesicle protein, synaptotagmin, may be linked to the calcium channels²², and the SNAREs and synaptotagmin may help mediate the calcium-triggered fusion of docked vesicles with the presynaptic membrane, thereby leading to the exocytosis of neurotransmitter during synaptic transmission^{28–30}. If the cytoplasmic domains of syntaxin and synaptotagmin are linked to the calcium channels at the frog's neuromuscular junction, they must pass through the AZM to reach the channels. Our observation that the ribs contact vesicles near the site where the

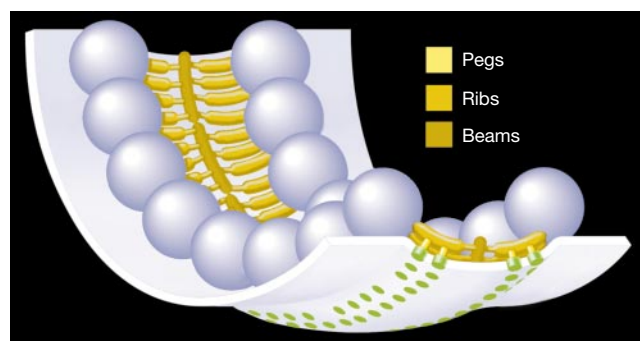


Figure 5 Arrangement and connections of pegs, ribs and beams.

vesicles are in close apposition to the presynaptic membrane raises the possibility that syntaxin and synaptotagmin reach the calcium channels through the ribs. The involvement of ribs in the fusion of vesicles with the presynaptic membrane would distinguish them from other AZM components connected to the vesicles, which may simply recognize and tether^{31,41} the vesicles.

In sum, we propose that the AZM is a multifunctional organelle that regulates the docking and fusion of synaptic vesicles at, and with, the presynaptic membrane, the anchoring of calcium channels in the membrane and the maintenance of spatial relationships between docked vesicles and channels. The material is also likely to be involved in the adhesion of the pre- to the postsynaptic membrane through the extracellular matrix^{5,32–34}. Electron microscope tomography studies aimed at determining how ribs and beams are associated with deeper components of the AZM, and how the material is associated with cytoplasmic components that mediate vesicle transport to the active zone, may indicate additional functions. Our general hypothesis should be testable at neuromuscular junctions and other synapses by combining EMT with well-established procedures for localizing specific proteins, and for studying the behaviour of synaptic organelles during synaptic transmission. Accurate information about the three-dimensional structural organization of the AZM and the molecular nature of its components is essential for a complete understanding of the presynaptic mechanisms that are involved in the regulation of synaptic transmission throughout the nervous system. □

Methods

Sample preparation

We used a conventional procedure^{23,35}. Briefly, cutaneous pectoris muscles of adult *Rana pipiens* (5 cm nose–rump length) were, in sequence, fixed with 1% phosphate buffered glutaraldehyde, refixed and stained with 1% osmium tetroxide in phosphate buffer, stained with saturated aqueous uranyl acetate, dehydrated in ethanol and propylene oxide, and embedded in Epon 812. Thin sections that had a silver interference colour (~50 nm thick) were mounted on 1 × 2 mm single slot Formvar coated grids, and stained with uranyl acetate, followed by lead citrate. For use in the alignment of tilt images, stained and mounted sections were floated on a drop of ~0.01% unconjugated gold colloid in H₂O (Sigma Chemical, St Louis, Missouri) for 1–2 min, and air dried.

Data collection

Data sets for MPI-9 and MPI-10 were made with a Philips CM200 electron microscope operated at 120 kV and equipped with a FEG, LN₂ cryoholder (Gatan, Warrendale, Pennsylvania) and 1024 × 1024 CCD (Tietz Video and Image Processing Systems GmbH, Gauting, Germany). The images were made on the CCD camera at a magnification of ×69,700, giving a single pixel size of 0.26 nm in the image. The tilt series was collected using an automatic procedure^{36,37}. Sections were pre-irradiated for 0.5–1.5 min, which, together with their low temperature, reduced shrinkage during data collection³⁸. For MPI-9, the tilt series consisted of 64 images spanning angles from –66° to +60° in 2° increments. For MPI-10, the series consisted of 67 images spanning angles from –62° to +70° in 2° increments. We made the data set for UC-1 on a Philips EM 430 electron microscope that was equipped for automatically collecting serial tilt images and was operated at 300 kV. It had a LaB₆ electron source, a cryoholder and a 1024 × 1024 CCD. The tilt series consisted of 140 images spanning angles from –70° to +69° in 1° increments. The images were made on the CCD camera at a magnification of ×30,600 with a bin size of two, so that the pixel size was 1.2 nm in the image.

Tomographic reconstruction

The gold beads on a section served as markers for aligning the tilt images. We used a scheme³⁵ that automatically detected the position of these markers and indexed them from image to image. The resulting set of position vectors was then arranged in a matrix representation of the tilt geometry, and the alignment information was obtained by inversion. The scheme provided an accuracy of < 1.3 pixels r.m.s. for MPI-9, MPI-10 and UC-1. We produced volume reconstructions by weighted backprojection³⁹. As each projection was combined into the volume, its grey-scale histogram was adjusted to compensate for uncontrolled variations in electron-beam exposure.

Signal restoration

We improved the signal-to-noise level in the reconstructed volume using nonlinear parametric signal estimation. A translation- and rotation-invariant tensor-wavelet transform was applied to serial two-dimensional volume slices perpendicular to the 0° tilt plane of the sample⁴⁰. Noise was reduced by discarding transform coefficients with values below a particular threshold. Image signal/noise ratio for the data presented here was improved by a factor of five, without significantly affecting the signal.

Segmentation

Structures were segmented from the reconstructed volume using a combination of manual and automatic methods applied to a series of two-dimensional slices. Each segmentation defined the boundaries of a volume of interest (VOI), a subset of the entire volume. We used two schemes to define the VOIs. For the presynaptic membrane and synaptic vesicles, which were heavily stained and had a simple geometry, a semi-automatic scheme was used in which the operator needed only to mark an anchor path on a single volume slice. We then automatically propagated this path perpendicular to the slice plane using a new, constrained grey-scale neighborhood-search scheme. For active zone material, which had a complex geometry and light to moderate stain, VOIs were defined by manually marking a closed path on a series of slices. The thickness of each slice for MPI-9 and MPI-10 was 0.52 nm (for example, Fig. 2c, d), and for UC-1 it was 1.2 nm. For all segmentation, the angular orientation of the slice plane was completely adjustable to maximize contrast boundary discrimination of the stained structure under study. The segmentation VOIs were slightly larger than the stained structures that they enclosed, to allow for accurate and complete isodensity surface calculations for surface renderings.

Surface rendering

To observe the stained structures within each VOI, a surface was calculated on the basis of a particular stain-density (grey) value. Initially, the isodensity value for each VOI was chosen to correspond to the 60% population value on the grey-scale cumulative distribution function; in some cases the value was subsequently adjusted slightly by the operator to correct for local inhomogeneities in stain density.

Received 14 September; accepted 24 November 2000.

- Peters, A., Palay, S. L. & Webster, H. de F. *The Fine Structure of the Nervous System* 198–203 (Oxford, Oxford, 1991).
- Katz, B. *The Release of Neural Transmitter Substances* (Thomas, Springfield, 1969).
- Couteaux, R. & Pécot-Dechavassine, M. Vésicules synaptiques et poches au niveau des "zones actives" de la jonction neuromusculaire. *Compt. Rend.* **271**, 2346–2349 (1970).
- Heuser, J. E. & Reese, T. S. Structural changes after transmitter release at the frog neuromuscular junction. *J. Cell Biol.* **88**, 564–580 (1981).
- Palay, S. L. Synapses in the central nervous system. *J. Biophys. Biochem. Cytol.* **2** (Suppl.), 193–202 (1956).
- Gray, E. G. Electron microscopy of presynaptic organelles of the spinal cord. *J. Anat.* **97**, 101–106 (1963).
- Gray, E. G. Problems of interpreting the fine structure of vertebrate and invertebrate synapses. *Int. Rev. Gen. Exp. Zool.* **2**, 139–170 (1966).
- Heuser, J. E. & Reese, T. S. in *Handbook of Physiology* Vol. 1, 261–294 (American Physiological Society, Bethesda, 1977).
- Frank, J. *Electron Tomography: Three-Dimensional Imaging with the TEM* (Plenum, New York, 1992).
- Radermacher, M. Three-dimensional reconstruction of single particles from random and nonrandom tilt series. *J. Electron Microsc. Tech.* **9**, 359–394 (1988).
- Zhou, Z., Prasad, B., Jakana, J., Rixon, F. & Chiu, W. Protein subunit structures in the herpes simplex virus capsid from 400 kV spot-scan electron cryomicroscopy. *J. Mol. Biol.* **242**, 456–469 (1994).
- Zwickl, P. et al. Primary structure of the thermoplasma proteasome and its implications for the structure, function, and evolution of the multicatalytic proteinase. *Biochemistry* **31**, 964–972 (1994).
- Moritz, M. et al. Three-dimensional structural characterization of centrosomes from early *Drosophila* embryos. *J. Cell Biol.* **130**, 1149–1159 (1995).
- Hoenger, A. & Aebi, U. 3-D reconstructions from ice-embedded and negatively stained biomacromolecular assemblies: A critical comparison. *J. Struct. Biol.* **117**, 99–116 (1996).
- McEwen, B. F., Arena, J. T., Frank, J. & Rieder, C. L. Structure of the colcemid-treated PtK1 kinetochore outer plate as determined by high voltage electron microscopic tomography. *J. Cell Biol.*

120, 301–312 (1993).

- Perkins, G. et al. Electron tomography of neuronal mitochondria: three-dimensional structure and organization of cristae and membrane contacts. *J. Struct. Biol.* **119**, 260–272 (1997).
- Ladinsky, M. S., Mastronarde, D. N., McIntosh, J. R., Howell, K. E. & Staehelin, L. A. Golgi structure in three dimensions: functional insights from the normal rat kidney cell. *J. Cell Biol.* **144**, 1135–1149 (1999).
- Lenzi, D., Runyeon, J. W., Crum, J., Ellisman, M. H. & Roberts, W. M. Synaptic vesicle populations in saccular hair cells reconstructed by electron tomography. *J. Neurosci.* **19**, 119–132 (1999).
- Heuser, J. E., Reese, T. S. & Landis, D. M. D. Functional changes in frog neuromuscular junctions studied with freeze-fracture. *J. Neurocytol.* **3**, 109–131 (1974).
- Robitaille, R., Alder, E. M. & Charlton, M. P. Strategic location of calcium channels at transmitter release sites of frog neuromuscular synapses. *Neuron* **5**, 773–779 (1990).
- Stanley, E. F. The calcium channel and the organization of the presynaptic transmitter release face. *Trends Neurosci.* **20**, 404–409 (1997).
- Catterall, W. A. Structure and function of neuronal Ca²⁺ channels and their role in neurotransmitter release. *Neuron* **24**, 307–323 (1998).
- McMahan, U. J. & Slater, C. R. The influence of basal lamina on the accumulation of acetylcholine receptors at synaptic sites in regenerating muscle. *J. Cell Biol.* **98**, 1453–1473 (1984).
- Colledge, M. & Froehner, S. C. To muster a cluster: anchoring neurotransmitter receptors at synapses. *Proc. Natl Acad. Sci. USA* **95**, 3341–3345 (1998).
- Robitaille, R., Garcia, M. L., Kaczorowski, G. J. & Charlton, M. P. Functional colocalization of calcium and calcium-gated potassium channels in control of transmitter release. *Neuron* **11**, 645–655 (1993).
- Cohen, M. W., Hoffstrom, B. G. & DeSimone, D. W. Active zones on motor nerve terminals contain $\alpha 3\beta 1$ integrin. *J. Neurosci.* **20**, 4912–4921 (2000).
- Weber, T. et al. SNAREpins: minimal machinery for membrane fusion. *Cell* **92**, 759–772 (1998).
- Hazuka, C. D., Foletti, D. L. & Scheller, R. H. in *Neurotransmitter Release* (ed. Bellen, H. J.) 81–125 (Oxford Univ. Press, 1999).
- Jahn, R. & Südhof, T. C. Membrane fusion and exocytosis. *Annu. Rev. Biochem.* **68**, 863–911 (1999).
- Broadie, K., Bellen, H. K., DiAntonio, J., Littleton, J. T. & Schwarz, T. L. Absence of synaptotagmin disrupts excitation-secretion coupling during synaptic transmission. *Proc. Natl Acad. Sci. USA* **91**, 10727–10731 (1994).
- Pfeffer, S. R. Transport-vesicle targeting: tethers before SNAREs. *Nature Cell Biol.* **1**, E17–E22 (1999).
- Sunderland, W. J., Son, Y. -J., Miner, J. H., Sanes, J. R. & Carlson, S. S. The presynaptic calcium channel is part of a transmembrane complex linking a synaptic laminin ($\alpha 4\beta 2\gamma 1$) with non-erythroid spectrin. *J. Neurosci.* **20**, 1009–1019 (2000).
- Martin, P. T., Kaufman, S. J., Kamer, R. H. & Sanes, J. R. Synaptic integrins in developing, adult, and mutant muscle: selective association of $\alpha 1$, $\alpha 7A$, $\alpha 7B$ integrins in the neuromuscular junction. *Dev. Biol.* **174**, 125–139 (1996).
- Bewick, G. S., Young, C. & Slater, C. R. Spatial relationships of utrophin, dystrophin, β -dystroglycan, and β -spectrin to acetylcholine receptor clusters during postnatal maturation of the rat neuromuscular junction. *J. Neurocytol.* **25**, 367–369 (1996).
- Ress, D., Harlow, M., Schwarz, M., Marshall, R. M. & McMahan, U. J. Automatic acquisition of fiducial markers and alignment of images in tilt series for electron tomography. *J. Electron Microsc.* **48**, 277–287 (1999).
- Dierksen, K., Typke, D., Hegerl, R. & Baumeister, W. Towards automatic electron tomography. *Ultramicroscopy* **49**, 109–120 (1992).
- Dierksen, K. et al. Three dimensional structure of lipid vesicles embedded in vitreous ice and investigated by automatic electron tomography. *J. Biophys.* **68**, 1416–1422 (1995).
- Braunfeld, M., Koster, A., Sedat, J. & Agard, D. Cryo automated electron tomography: towards high-resolution reconstructions of plastic-embedded structures. *J. Microsc.* **174**, 75–84 (1994).
- Radermacher, M. in *Electron Tomography: Three-Dimensional Imaging with the TEM* (ed. Frank, J.) 91–115 (Plenum, New York, 1992).
- Stoschek, A. & Hegerl, R. Denoising of electron tomographic reconstructions using multiscale transformations. *J. Struct. Biol.* **120**, 257–265 (1997).
- Wang, X. et al. Aczonin, a 550-kD putative scaffolding protein of presynaptic active zones, shares homology regions with Rim and Bassoon and binds profilin. *J. Cell Biol.* **147**, 151–162 (1999).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We are grateful to W. Baumeister and members of the Department of Structural Biology at the Max-Planck Institute for Biochemistry in Martinsried, Germany, for use of their CM 200 electron microscope; A. Koster participated in the data collection for MPI-9 and MPI-10. We are also grateful to D. Agard and members of his laboratory in the Department of Biochemistry and Biophysics at the University of California San Francisco for use of the Philips EM 430; M. Braunfeld participated in the data collection for UC-1. T. Schwarz at Harvard Medical School provided comments on the manuscript, and B. Colyear at Stanford University made the schematic illustrations. This work was supported by grants from the NIH and the Deutsche Forschungsgemeinschaft, and by the M. A. Antelman Fund.

Correspondence and requests for materials should be addressed to U.J.M. (e-mail: ujmcmahan@stanford.edu).

Dynamos in asymptotic-giant-branch stars as the origin of magnetic fields shaping planetary nebulae

Eric G. Blackman*, Adam Frank*, J. Andrew Markiel†, John H. Thomas*‡ & Hugh M. Van Horn*

* Department of Physics and Astronomy and C. E. K. Mees Observatory, University of Rochester, Rochester, New York 14627-0171, USA

† Department of Astronomy, University of Washington, Seattle, Washington 98195-1580, USA

‡ Department of Mechanical Engineering, University of Rochester, Rochester, New York 14627-0171, USA

Planetary nebulae are thought to be formed when a slow wind from the progenitor giant star is overtaken by a subsequent fast wind generated as the star enters its white dwarf stage¹. A shock forms near the boundary between the winds, creating the relatively dense shell characteristic of a planetary nebula. A spherically symmetric wind will produce a spherically symmetric shell, yet over half of known planetary nebulae are not spherical; rather, they are elliptical or bipolar in shape². A magnetic field could launch and collimate a bipolar outflow, but the origin of such a field has hitherto been unclear, and some previous work has even suggested that a field could not be generated³. Here we show that an asymptotic-giant-branch (AGB) star can indeed generate a strong magnetic field, having as its origin a dynamo at the interface between the rapidly rotating core and the more slowly rotating envelope of the star. The fields are strong enough to shape the bipolar outflows that produce the observed bipolar planetary nebulae. Magnetic braking of the stellar core during this process may also explain the puzzlingly⁴ slow rotation of most white dwarf stars.

One model for producing bipolar or elliptical planetary nebulae assumes that the slow wind from the progenitor star is denser near the equatorial plane than along the poles^{5–8}. Such an asymmetry can result if the progenitor is part of a close binary star system. However, the incidence of close binaries seems not to be sufficient to account for all asymmetric planetary nebulae.

Magnetic shaping is a promising mechanism for forming asymmetric planetary nebulae. A toroidal magnetic field can constrain the flow in the equatorial plane⁹, redirecting it preferentially along the poles¹⁰, and a dipole field can create a dense torus around the central star which then collimates the wind¹¹. The strength of a toroidal magnetic field in the wind will be increased in the shocked bubble, leading to collimation^{12,13}. Collimation might also occur close to the star through magneto-centrifugal processes¹⁴.

We suggest that the required magnetic field is generated by a dynamo in the central star during its AGB phase. The necessary ingredients for an alpha–omega interface dynamo are very likely to be present near the outer boundary of the degenerate core of an AGB star. Contraction of the core and expansion of the envelope of the star during this stage of evolution will tend to produce strong differential rotation (the omega effect) regardless of the initial rotation profile of the star on the main sequence. In the lower part of the deep, rotating convection zone the necessary alpha effect (which produces poloidal field from toroidal) is provided by the helical convection itself or by magnetic instabilities. We note that a previous argument against significant dynamo action in an AGB star³ was based on a model of a distributed dynamo operating solely within the slowly rotating envelope, a very different picture from the one we present here.

To investigate the possibility of dynamo action in AGB stars, we first estimate the rotation profile in a typical AGB star evolving from

a rotating main-sequence star, and we then use this rotation profile in a self-consistent nonlinear dynamo model to test for dynamo action. For a typical AGB star, we use an evolutionary model (S. D. Kawaler, personal communication) for a star of mass $3M_{\odot}$, where $M_{\odot}$ is the mass of the Sun. We assume that this star is rotating uniformly on the main sequence with angular velocity $1 \times 10^{-4} \text{ s}^{-1}$ (corresponding to a surface rotation velocity of $\sim 200 \text{ km s}^{-1}$; ref. 15) and that during the subsequent evolution the angular momentum of each spherical mass shell is conserved. (As we do expect some angular momentum exchange to occur in any real star, for example by magnetic braking, our model exaggerates the degree of differential rotation that may develop; nevertheless, we expect our model to be at least qualitatively correct.) The resulting rotation profile for this star at the tip of the AGB (Kawaler's model 1401, hereafter SDK 1401) has strong differential rotation, with the innermost core rotating faster and the outer envelope rotating much slower than the initial main-sequence rotation rate. The rotation profile in the neighbourhood of the core–envelope interface is shown in Fig. 1. Our resulting rotation profile and thus our approach are strongly supported by observations that have been interpreted to imply core–envelope decoupling: observed fast rotation in horizontal-branch stars^{16–18} and the existence of young, rapidly spinning pulsars¹⁹.

The arrangement of differential rotation and convection in adjacent layers in this star is similar to that in the Sun and strongly suggests the possibility of an alpha–omega interface dynamo. To test this possibility we use a simple nonlinear alpha–omega dynamo model that we have used previously to study dynamos in white dwarf stars²⁰. This model, based on a local analysis of the full mean-field dynamo equations, assumes that the dynamo region consists of two adjacent, thin layers—an inner layer in differential rotation, and an outer layer of convection in which the alpha effect operates. The solar dynamo is thought to be of this type^{21,22}, and this configuration also corresponds roughly to the situation in our model AGB star. The nonlinear saturation of oscillatory dynamo modes in our model is caused by quenching of the alpha effect by the generated magnetic field. The free parameters in the model are calibrated by applying the model to the Sun and requiring it to produce a dynamo with the correct period (22 yr) and appropriate amplitude (toroidal field strengths of $\sim 10^4 \text{ G}$). The calibrated model is then applied to the model AGB star. As input to our

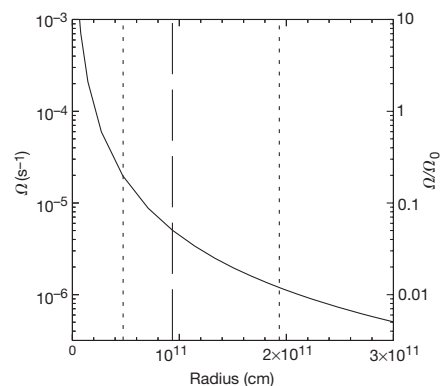


Figure 1 Internal rotation rate of our model $3.0M_{\odot}$ asymptotic-giant-branch star as a function of radius near the base of the convection zone. The angular velocity Ω is given both in units of s^{-1} and in units of the initial (uniform) rotation rate Ω_0 of the star when on the main sequence. The inner core has been spun up by contraction, whereas the outer layers rotate at a much slower rate than the initial rate, owing to the large expansion of the envelope. The dashed line indicates the location of the base of the convection zone; the inner dotted line indicates the inner boundary of the differential-rotation layer assumed in our dynamo model; and the outer dotted line indicates the outer boundary of the alpha-effect layer.

dynamo model, we obtain the following values from our model of differential rotation in SDK 1401: a differential rotation layer of thickness 4.6×10^{10} cm lying just below the convection zone (at radius 9.35×10^{10} cm), with angular velocity dropping from $2 \times 10^{-5} \text{ s}^{-1}$ to $5 \times 10^{-6} \text{ s}^{-1}$ across this layer; and an alpha-effect layer of thickness 1×10^{11} cm lying just above the base of the convection zone, with typical convective velocity $1 \times 10^5 \text{ cm s}^{-1}$ and density $6.6 \times 10^{-4} \text{ g cm}^{-3}$ (see Fig. 1).

For these input values for SDK 1401 we obtain an oscillatory dynamo with a period of ~ 0.4 yr and a maximum mean toroidal field strength of $\sim 5 \times 10^4$ G just below the convection zone. (This result is robust, in the sense that it is not sensitive to moderate changes in the values of the input parameters.) The ratio of thermal pressure to magnetic pressure just below the convection zone is $\beta \approx 10^3$. The field need not be volume-filling, and flux tube formation is very likely as the magnetic field becomes subject to buoyancy instabilities. The fraction of volume filled by flux tubes, and thus the surface area covering fraction f for long thin tubes, could be as low as β^{-1} (ref. 23). The strength of a typical flux tube in this case, determined by pressure balance across the tube, is 3×10^6 G. The strength of such a tube upon rising to the top of the convection zone ($R = 5.5 \times 10^{12}$ cm) would be ~ 400 G in pressure balance. Surface magnetic fields of this strength could be associated with flares and coronal mass ejections from the AGB star by analogy with the Sun²⁴. These coronal mass ejections could be responsible for the non-axisymmetric, collimated shapes with unpaired ejections of knots or bullets seen in some planetary nebulae^{25,26}.

An average toroidal field strength of 5×10^4 G at the interface represents 10^{41} erg of magnetic energy in the volume of the shear layer. The rate of magnetic energy flow from the interface layer due to buoyancy is about $(B^2/8\pi)(4\pi R_c^2)v_A$, where v_A is the Alfvén speed in the layer and R_c is the radial location of the layer, at the outer boundary of the core. For $B = 5 \times 10^4$ G, a density in the interface layer of $1.3 \times 10^{-3} \text{ g cm}^{-3}$, and $R_c = 9.35 \times 10^{10}$ cm, this gives an upward magnetic energy supply rate of $\sim 4 \times 10^{36} \text{ erg s}^{-1}$. This number is comparable to the turbulent energy dissipation rate in the convection zone, given by $(\rho v^2)(\nu/l)(V_{\text{conv}}) \approx 10^{36} \text{ erg s}^{-1}$, when estimated for typical values, such as, $\rho \approx 10^{-4} \text{ g cm}^{-3}$ for the density, $V_{\text{conv}} \approx 5 \times 10^3 \text{ cm}^3$ for the volume, $\nu \approx 10^5 \text{ cm}^2 \text{ s}^{-1}$ for the turbulent speed, and $l \approx 10^{11}$ cm for the characteristic eddy scale. This means that the buoyancy-driven magnetic energy supply rate is an upper limit on the supply rate to the corona, since a sizeable fraction may be shredded and dissipated on its way up through the convection zone. The fraction that does make it to the corona would be available for particle acceleration and coronal emission from the AGB star. Some of this could take the form of localized flares and coronal mass ejections, as on the Sun.

Detection of X-rays from such a corona is unlikely while the star is in either the AGB stage or the proto-planetary-nebula stage owing to the high density of the wind. We estimate the optical depth for soft (1 keV) and hard (10 keV) X-rays to be 10 and 10^3 , respectively. The X-rays could also produce a layer of highly ionized atoms near the star, but this would probably be enshrouded by much cooler gas in the stellar wind and thus be hard to detect. The deposition of energy into the corona may have consequences for the thermal and dynamical structure of the winds, which will need further investigation. We expect that the corona will be removed along with the bulk of the magnetic flux by the time the star reaches the mature planetary nebula phase.

The core of an AGB star can be expected to rotate very rapidly because of contraction, as illustrated here by our model. However, the white dwarf stars, which develop from AGB cores following the planetary nebula phase, generally rotate much more slowly than expected if angular momentum conservation of the AGB core is assumed. This well known problem⁴ might be resolved by invoking the dynamo action proposed here; the strong magnetic field

generated by the AGB dynamo could produce magnetic braking. This would occur during and after the time the AGB star sheds its envelope. Because of flux freezing, the field lines are drawn out with the envelope material but are anchored in the core. In this way the angular momentum of the core is transported outward to the envelope, and the core is spun down by a post-AGB magnetohydrodynamic (MHD) wind¹⁴. During the braking period, bipolar and multipolar MHD winds could form the shapes observed in some planetary nebulae as the stellar convective envelope is shed.

The MHD wind luminosity then gives an upper limit to the kinetic power¹⁴. For an MHD wind from field lines anchored in the core, the luminosity can be as high as $L_{\text{MHD}} \approx 2 \times 10^{38} (B/10^5 \text{ G})^2 (\Omega_c/2 \times 10^{-5} \text{ s}^{-1}) (R_c/10^{11} \text{ cm})^3 \text{ erg s}^{-1}$, where Ω_c is the angular speed of the core material where the field lines are anchored, and R_c is the core radius. An approximately 5% conversion of this is required to account for the energy loss rates of $10^{37} (\dot{M}/6 \times 10^{21} \text{ g s}^{-1}) (V/400 \text{ km s}^{-1})^2 \text{ erg s}^{-1}$ characteristic of some proto-planetary nebulae²⁷, where $\dot{M}$ is the mass loss rate and V is the wind speed. The timescale for the MHD wind spin-down of the post-AGB star is then $\tau_s = 140 (M_c/M_\odot) (\Omega_c/2 \times 10^{-5} \text{ s}^{-1}) (B/10^4 \text{ G})^{-2} (R_c/10^{11} \text{ cm})^{-1} \text{ yr}$.

The post-AGB wind, produced when the AGB star sheds its outer layers and exposes the rapidly rotating, magnetized core, may be strongly collimated by magneto-centrifugal processes²⁸. The degree to which a magnetized rotator will collimate a wind driven off its surface can be expressed by a rotation parameter²⁹ Q , given by $Q/Q_\odot \approx 4(\psi_c/5 \times 10^{26} \text{ G cm}^2) (\Omega_c/2 \times 10^{-5} \text{ s}^{-1}) \times (\dot{M}/6 \times 10^{21} \text{ g s}^{-1})^{-1/2} (V/400 \text{ km s}^{-1})^{-3/2}$, where ψ_c is the magnetic flux at large distances (proportional to BR^2), V is the outflow speed, and $Q_\odot \approx 0.12$ is the value for the solar wind (using $V_\odot = 400 \text{ km s}^{-1}$, $\dot{M}_\odot = 1.6 \times 10^{12} \text{ g s}^{-1}$, $\Omega_\odot = 3 \times 10^{-6} \text{ s}^{-1}$, and $\psi_\odot = 1.4 \times 10^{22} \text{ G cm}^2$). We have scaled the flux to the upper limit using $B \approx 5 \times 10^4$ G and $R_c \approx 10^{11}$ cm for our dynamo-produced field strength at the AGB core, and we have scaled the outflow parameters $\dot{M}$ and V using proto-planetary nebulae values²⁷. For $Q \geq 1$ the system is classified as a fast magnetic rotator. The larger Q is, the more strongly self-collimated the outflow is³⁰. We can see that for these parameters collimation greater than that in the solar wind is possible. If, instead, we use representative outflow parameters for later stages of planetary nebulae³¹, $\dot{M} = 5 \times 10^{-7} M_\odot \text{ yr}^{-1}$ and $V = 1,000 \text{ km s}^{-1}$, we find $Q/Q_\odot \approx 15$, which implies that significant collimation is possible. Magnetic collimation may thus be intrinsic to these sources and may not require shocks, as in some models¹³.

Our model leads us to predict that magnetic fields should be apparent in the winds of AGB stars and proto-planetary nebulae. Such magnetic fields should be detectable in some maser spots, and indeed have already been observed in at least one object³² with field strengths consistent with our model (based on flux conservation), although this field might be locally amplified by turbulence. We also predicted that strongly collimated flows in proto-planetary nebulae should have signatures of ordered magnetic fields (for example, polarization, Faraday rotation), reflecting the role of the magnetic field in their launching and collimating. At small distances from the central star, the field should be primarily poloidal and parallel to the jet flow, whereas at large distances we would expect a dominant toroidal component (perpendicular to the outflow) as the hoop stresses take over the collimation.

Our dynamo model depends on rapid rotation of the AGB core. If a way can be found to measure the rotation rates of AGB cores, it could serve as an observational test of our model. Stars that rotate much more slowly than average on the main sequence would not be expected to produce significant dynamos at the AGB stage and hence would not be expected to produce collimated outflows by magnetic shaping. On the other hand, we would expect those AGB stars that do have rapidly rotating cores to produce bipolar planetary nebulae.

As the rotation rate of the degenerate core is reduced by magnetic braking and the convective envelope is removed, the stellar dynamo will shut down. Some remnant field anchored in the core will survive even without a convection zone, although the convective envelope may not be removed completely. Indeed, white dwarfs do have thin surface convection zones which can support a near-surface dynamo in the white dwarf itself²⁰.

We have thus demonstrated that dynamos are likely to operate in AGB stars. As a star evolves off the AGB, the dynamo-generated magnetic field will be strong enough to drive a strong, self-collimating outflow and to slow the rotation of the core by magnetic braking. Eruptions analogous to coronal mass ejections, expected as a consequence of the dynamo activity, could produce asymmetric structures in the wind. Thus our model opens up the possibility of constructing a new, self-consistent paradigm for planetary-nebula formation, beginning with AGB stars and ending with slowly rotating white dwarfs. □

Received 10 August; accepted 28 November 2000.

1. Kwok, S., Purton, C. R. & Fitzgerald, P. M. On the origin of planetary nebulae. *Astrophys. J.* **219**, L125–L127 (1978).
2. Manchado, A., Villaver, E., Stanghellini, L. & Guerrero, M. in *Asymmetrical Planetary Nebulae II. From Origins to Microstructures* (eds Kastner, J. H., Stoker, N. & Rappaport, S.) 17–23 (ASP Conf. Ser. 199, Astronomical Society of the Pacific, San Francisco, 2000).
3. Soker, N. & Harpaz, A. Can a single AGB star form an axially symmetric planetary nebula? *J. Astron. Soc. Pac.* **104**, 923–930 (1992).
4. Koester, D., Dreizler, S., Weidemann, V. & Allard, N. F. Search for rotation in white dwarfs. *Astron. Astrophys.* **338**, 612–622 (1998).
5. Balick, B. The evolution of planetary nebulae. I—Structures, ionizations, and morphological sequences. *Astron. J.* **94**, 671–678 (1987).
6. Soker, N. & Livio, M. Interacting winds and the shaping of planetary nebulae. *Astrophys. J.* **339**, 268–278 (1989).
7. Mellema, G., Eulerink, F. & Icke, V. Hydrodynamical models of aspherical planetary nebulae. *Astron. Astrophys.* **252**, 718–732 (1991).
8. Icke, V., Balick, B. & Frank, A. The hydrodynamics of aspherical planetary nebulae. II. Numerical modelling of the early evolution. *Astron. Astrophys.* **253**, 224–243 (1992).
9. Gurzadyan, G. *Planetary Nebulae* (Gordon and Breach, New York, 1969).
10. Pascoli, G. La nature des nébuleuses planétaires bipolaires. *Astron. Astrophys.* **180**, 191–200 (1987).
11. Matt, S., Balick, B., Winglee, R. & Goodson, A. Disk formation by asymptotic giant branch winds in dipole magnetic fields. *Astrophys. J.* **545**, 965–973 (2000).
12. Chevalier, R. A. & Luo, D. Magnetic shaping of planetary nebulae and other stellar wind bubbles. *Astrophys. J.* **421**, 225–235 (1994).
13. García-Segura, G., Langer, N., Rozyczka, M. & Franco, J. Shaping bipolar and elliptical planetary nebulae: effects of stellar rotation, photoionization heating, and magnetic fields. *Astrophys. J.* **517**, 767–781 (1999).
14. Blackman, E. G., Frank, A. & Welch, C. MHD stellar and disk winds: Application to planetary nebulae. *Astrophys. J.* (in the press).
15. Cox, A. N. *Allen's Astrophysical Quantities* 4th edn, 389 (AIP Press, New York, 1999).
16. Peterson, R. C. The rotation of horizontal-branch stars. II. Members of the globular clusters M3, M5, and M13. *Astrophys. J.* **275**, 737–751 (1983).
17. Pinsonneault, M., Deliyannis, C. P. & Demarque, P. Evolutionary models of halo stars with rotation. I. Evidence for differential rotation with depth in stars. *Astrophys. J.* **367**, 239–252 (1991).
18. Behr, B. B. et al. A new spin on the problem of horizontal-branch gaps: stellar rotation along the blue horizontal branch of globular cluster M13. *Astrophys. J.* **528**, 849–853 (2000).
19. Livio, M. & Pringle, J. E. The rotation rates of white dwarfs and pulsars. *Astrophys. J.* **505**, 339–343 (1998).
20. Thomas, J. H., Markiel, J. A. & Van Horn, H. M. Dynamo generation of magnetic fields in white dwarfs. *Astrophys. J.* **453**, 403–410 (1995).
21. Parker, E. N. A solar dynamo surface wave at the interface between convection and nonuniform rotation. *Astrophys. J.* **408**, 707–719 (1993).
22. Markiel, J. A. & Thomas, J. H. Solar interface dynamo models with a realistic rotation profile. *Astrophys. J.* **523**, 827–837 (1999).
23. Blackman, E. G. Overcoming the back reaction on turbulent motions in the presence of magnetic fields. *Phys. Rev. Lett.* **77**, 2694–2697 (1996).
24. Pick, M. in *Ninth European Meeting on Solar Physics: Magnetic Fields and Solar Processes* (ed. Wilson, A.) 183 (ESA SP Series SP-448, European Space Agency, 1999).
25. Trammell, S. R. in *Asymmetrical Planetary Nebulae II. From Origins to Microstructures* (eds Kastner, J. H., Soker, N. & Rappaport, S.) 147–150 (ASP Conf. Ser. 199, Astronomical Society of the Pacific, San Francisco, 2000).
26. Sahai, J. in *Asymmetrical Planetary Nebulae II. From origins to Microstructures* (eds Kastner, J. H., Soker, N. & Rappaport, S.) 209–216 (ASP Conf. Ser. 199, Astronomical Society of the Pacific, San Francisco, 2000).
27. Alcolea, J. et al. in *Asymmetrical Planetary Nebulae II. From Origins to Microstructures* (eds Kastner, J. H., Soker, N. & Rappaport, S.) 347–354 (ASP Conf. Ser. 199, Astronomical Society of the Pacific, San Francisco, 2000).
28. Königl, A. & Pudritz, R. in *Protostars and Planets IV* (eds Mannings, V., Boss, V. A. & Russell, S.) 759–791 (Univ. Arizona Press, Tucson, 2000).
29. Tsinganos, K. & Bogovalov, S. Magnetic collimation of solar and stellar winds. *Astron. Astrophys.* **356**, 989–1002 (2000).

30. Lery, T., Heyvaerts, J., Appl, S. & Norman, C. Outflows from magnetic rotators. *Astron. Astrophys.* **337**, 603–624 (1998).
31. Balick, B. et al. FLIERs and other microstructures in planetary nebulae. IV. Images of elliptical PNs from the Hubble Space Telescope. *Astron. J.* **116**, 360–371 (1999).
32. Palen, S. & Fix, J. D. Models of OH maser variations in U Herculis. *Astrophys. J.* **531**, 391–400 (2000).

Acknowledgements

We are grateful to S. Kawaler and to F. D'Antona and P. Ventura for making available to us detailed tables of their evolutionary models for AGB stars. This work was supported by the NSF, NASA and the DOE.

Correspondence and requests for materials should be addressed to J.H.T.
(e-mail: thomas@astro.me.rochester.edu).

Geochemical evidence for magmatic water within Mars from pyroxenes in the Shergotty meteorite

Harry Y. McSween Jr^{*}, Timothy L. Grove[†], Rachel C. F. Lentz^{*}, Jesse C. Dann[‡], Astrid H. Holzheid[‡], Lee R. Riciputi[§] & Jeffrey G. Ryan[§]

^{*} Department of Geological Sciences, University of Tennessee, Knoxville, Tennessee 37996, USA

[†] Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

[‡] Chemical and Analytical Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA

[§] Department of Geology, University of South Florida, Tampa, Florida 33620, USA

Observations of martian surface morphology have been used to argue that an ancient ocean once existed on Mars¹. It has been thought that significant quantities of such water could have been supplied to the martian surface through volcanic outgassing, but this suggestion is contradicted by the low magmatic water content that is generally inferred from chemical analyses of igneous martian meteorites². Here, however, we report the distributions of trace elements within pyroxenes of the Shergotty meteorite—a basalt body ejected 175 million years ago from Mars³—as well as hydrous and anhydrous crystallization experiments that, together, imply that water contents of pre-eruptive magma on Mars could have been up to 1.8%. We found that in the Shergotty meteorite, the inner cores of pyroxene minerals (which formed at depth in the martian crust) are enriched in soluble trace elements when compared to the outer rims (which crystallized on or near to the martian surface). This implies that water was present in pyroxenes at depth but was largely lost as pyroxenes were carried to the surface during magma ascent. We conclude that ascending magmas possibly delivered significant quantities of water to the martian surface in recent times, reconciling geologic and petrologic constraints on the outgassing history of Mars.

The Shergotty meteorite is a martian basalt (shergottite), approximately 175 million years old³, with measured water contents⁴ of only 130–350 parts per million (p.p.m.). Pyroxenes (pigeonite and augite) were the earliest crystallizing minerals, and the homogeneous magnesian cores of these grains are thought to have formed at depth and been entrained in the ascending magma^{3,5}. During ascent or eruption, the pyroxenes developed iron-rich rims, as plagioclase and other minerals joined the crystallization sequence. The pyroxene core and rim compositions thus reflect magma compositions at depth and near (or on) the planet's surface,

respectively. The presence of hydrous amphibole within trapped melt inclusions in pyroxene cores previously fuelled speculation that the Shergotty magma contained 1–2 wt% water at depth⁶, but conflicting interpretations of the low measured hydrogen contents in these amphiboles^{7–9} suggest that this evidence is inconclusive.

Abundances of trace elements in cores and rims also record the magma's evolution. Analyses of Li, Be, B, Ce, Y and Ti in Shergotty pyroxenes using secondary-ion mass spectrometry are summarized in Table 1 of Supplementary Information. During fractional crystallization of pyroxenes, all these incompatible elements should partition into the melt, and thus their abundances should increase from pyroxene cores to rims. As expected, Be and Ti increase from core to rim, whereas B and Li decrease (Fig. 1a, b). Ce also appears to decrease relative to Y, which behaves similarly to heavy rare-earth elements (Fig. 1c).

The presence of hot aqueous fluid could alter the expected partitioning behaviour during fractional crystallization, because B and Li are soluble at temperatures greater than 350 °C (ref. 10). Partitioning of these elements between clinopyroxene, magma and fluids has been studied experimentally at magmatic temperatures¹¹, revealing clear preferences for the aqueous phase. Unlike most other rare-earth elements, cerium can exist in more than one valence state, but any Ce^{4+} in magma should be reduced by Fe^{2+} . Ce^{3+} has been suggested to be more soluble than other rare earths, possibly accounting for negative Ce anomalies observed in some subduction-zone lavas¹². Thus Ce might also partition into an aqueous fluid, although its concentration in pyroxene rims might also be affected if whitlockite co-crystallized. The observed depletions of B, Li, and perhaps Ce/Y in Shergotty pyroxene rims are

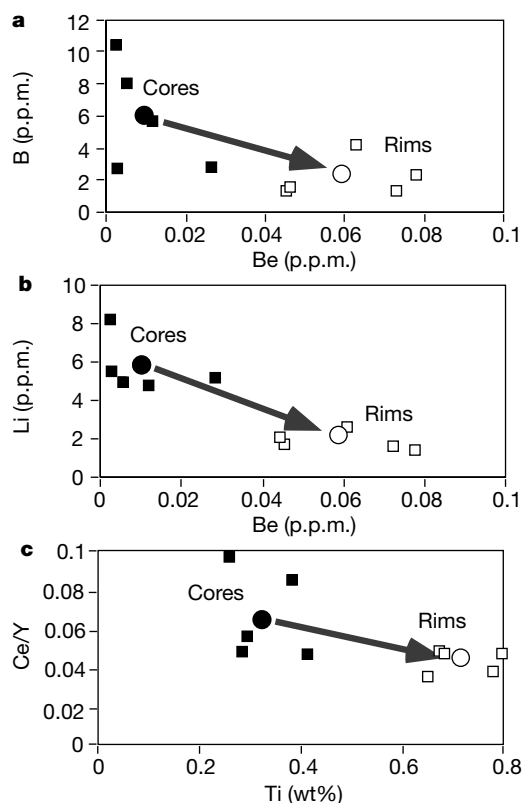


Figure 1 Measured abundances of B (a) and Li (b), and the measured Ce/Y ratio (c), relative to other incompatible elements (Be and Ti), in Shergotty pyroxenes. We interpret decreases in B, Li and Ce/Y from cores (filled symbols) to rims (open symbols) to reflect loss of soluble trace elements in an exsolved aqueous fluid. Analytical uncertainties for individual analyses (squares) are within the sizes of the points, and average values are indicated by circles.

consistent with removal of aqueous fluid after crystallization of the cores at depth. These observations suggest that water, carrying soluble trace elements, was exsolved and lost from the magma on ascent and eruption.

The Mg/(Mg+Fe) ratios of pigeonite and augite cores indicate that both pyroxenes in Shergotty crystallized simultaneously^{5,13}. Imaging shows that Shergotty contains approximately 13 vol.% of accumulated pyroxene cores, and the composition of crystal-free liquid has been estimated by subtracting pyroxenes from the bulk meteorite composition¹³. We performed crystallization experiments on this intercumulus liquid composition to determine if it would co-crystallize augite and pigeonite with compositions like those in Shergotty. Initial experimental conditions were anhydrous, at a pressure of 0.1 MPa and an oxygen fugacity set by the quartz–fayalite–magnetite (QFM) buffer. Compositions of pyroxenes crystallized under these conditions are summarized in Fig. 2. The Shergotty liquid composition did not produce a crystallizing assemblage at 0.1 MPa resembling the mineral assemblage in Shergotty. Instead of augite and pigeonite coexisting as liquidus phases, our 0.1-MPa results show a large crystallization interval for pigeonite (~100 °C) before an Fe-enriched low-wollastonite (low-Wo) augite appears (Fig. 2). Similar experimental results have been documented by other workers^{14,15}. Calculations of the crystallization sequence at elevated pressures¹³ and crystallization experiments undertaken at varying oxygen fugacities¹⁵ indicate that this discrepancy is not an effect of increased pressure or a different oxidation state.

We performed a second series of experiments at elevated water pressures (100 and 200 MPa) to test the influence of H_2O on the Shergotty crystallization sequence. At 100 MPa pigeonite was again the liquidus phase, and crystallized over a significant temperature interval (~80 °C) before augite appeared. The effect of the elevated H_2O content of the 200-MPa experiment was to expand the olivine primary phase volume so that the Shergotty intercumulus liquid was saturated near its liquidus with olivine, pigeonite and augite. Because the pigeonite and augite coexist as near-liquidus phases, pyroxenes from the 200-MPa experiments are very similar to the Shergotty pyroxene cores (Fig. 2). However, the CaSiO_3 (Wo) contents of the pigeonite cores in Shergotty are higher by ~4 mol% (average Wo content, 13 mol%) than that produced in the 200-MPa experiment (average Wo, 8.7 mol%) (see Table 2 of Supplementary Information). The difference in Wo content implies that pigeonite and augite in the Shergotty meteorite crystallized at

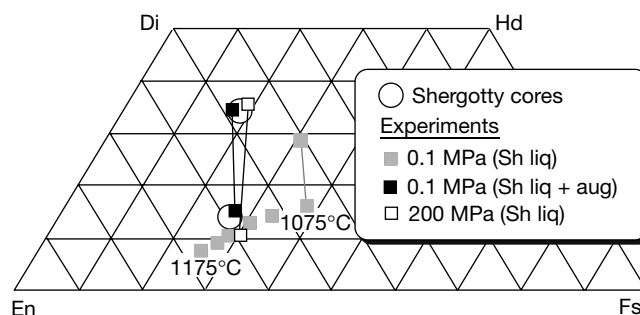


Figure 2 Comparison of the average compositions of Shergotty pigeonite and augite cores with experimental pyroxenes. The figure shows compositions of pyroxenes from anhydrous 0.1-MPa and water-saturated 200-MPa experiments, using the Shergotty liquid composition estimated in ref. 14 and a second composition which is an 85:15 mixture of this liquid and clinopyroxene (2 hedenbergite: 1 salite) with the same Mg/(Mg+Fe) ratio. Vertical bars connect the compositions of the first crystallizing coexisting pyroxenes. En, enstatite MgSiO_3 ; Fs, ferrosilite FeSiO_3 ; Di, diopside CaMgSiO_6 ; Hd, hedenburgite CaFeSiO_6 ; Sh liq, Shergotty intercumulus liquid; Sh liq+aug, Sh liq with augite added.

higher temperature than in our 200-MPa experiment. The effect of H_2O is to bring pigeonite and augite nearer the liquidus, but the 200-MPa experimental pyroxenes are not identical to the Shergotty pigeonite and augite cores.

To produce a melt saturated with two pyroxenes on its liquidus at lower water contents, the bulk composition must be modified slightly. We synthesized a new liquid composition by adding 15 wt% augite that had an $Mg/(Mg+Fe)$ ratio similar to that of the intercumulus liquid. This modified composition was run at 0.1 and 100 MPa. At 100 MPa, augite and pigeonite are liquidus phases at 1,150 °C. At 0.1 MPa, augite is the liquidus phase at 1,170 °C, and both augite and pigeonite are present at 1,150 °C at the QFM buffer. The average pigeonite composition in the 0.1-MPa run has 14.7 mol% Wo (Fig. 2). Therefore, the composition of the liquidus pigeonite and augite in Shergotty records a temperature and H_2O content that is between our 0.1-MPa anhydrous result and the hydrous 100- and 200-MPa experiments.

The effect of variations in temperature on the compositions of coexisting augite and pigeonite has been investigated previously¹⁶. The reported relation between temperature and the Wo content of pigeonite coexisting with augite on the pigeonite–augite solvus is identical to that obtained in the present work from the 0.1-MPa and 200-MPa Shergotty experiments. Therefore, the conditions of crystallization of the Shergotty pyroxenes can be used to infer the liquidus temperature and pre-eruptive H_2O pressure. The effects of variations in H_2O pressure on the temperature of the vapour-saturated liquidus are well known for natural basaltic magmas^{17,18}. Using the 0.1- and 200-MPa pigeonites that coexist with augite as a guide, we estimated by interpolation the liquidus temperature for augite and pigeonite in the Shergotty meteorite. We obtained a temperature of 1,120 °C and a magmatic H_2O content of ~1.8 wt%. These pre-eruptive conditions correspond to a pressure of ~55 MPa

using the model of ref. 19, or a depth of ~5 km in the martian crust. This pre-eruptive H_2O content is similar to the dissolved water content inferred for other martian magmas²⁰, and also to the composition of andesitic rocks determined by the Mars Pathfinder²¹.

The magmatic evolution of the Shergotty parent magma that we propose is illustrated in Fig. 3. Vesiculation and loss of water occurred after pyroxene cores crystallized at a depth of greater than 5 km. Although no vesicles are observed in Shergotty, a model of volatile exsolution in martian magmas²² indicates that outgassing should be gradual but continuous above this depth. Because of lower martian gravity, gas exsolution begins at greater depth than on Earth, and the low atmospheric pressure on Mars encourages thorough degassing as the magma ascends²³. Magma outgassing, even on Earth, can be very efficient; for example, viscous rhyolitic magma at Mono Craters, California, was found to lose more than 95% of its total water—from 2.4 to 0.11 wt%—on ascent²⁴.

Can we generalize the Shergotty meteorite results to Mars? Based on similarities in visible and near-infrared reflectance spectra, Shergotty-like basalts have been suggested to dominate large areas of the martian surface²⁵, although basalts mapped by thermal emission spectroscopy on the Mars Global Surveyor orbiter²⁶ appear to have spectra distinct from shergottites.

Although the water content of the Shergotty magma may have increased during fractionation before the crystallization of pyroxene phenocrysts, the presence of ~1.8 wt% water appears to be inconsistent with the dry mantle of the conventional Mars model². This model² estimates the abundance of water in the martian mantle to be 36 p.p.m., on the basis of the abundance of Cl and the relative solubilities of Cl and H_2O in basaltic magma. The same procedure would give an incorrect estimation of water in the Earth's mantle, because Cl has been sequestered in the crust. A newer Mars compositional model²⁷, based on matching oxygen isotopes, suggests a bulk-planet Cl abundance eight times higher than that estimated from element ratios, which would imply a correspondingly greater mantle water content. A value of several hundred p.p.m. of water is comparable to the terrestrial mantle; it is also consistent with observed enrichments of other volatile elements in martian meteorites², and allows higher magmatic water contents at reasonable degrees of partial melting. Alternatively, a Shergotty magma produced by anhydrous mantle melting could have interacted with metasomatized, hydrated mantle lithosphere during ascent.

It is also possible that the water in the Shergotty meteorite was added during assimilation of martian crustal materials or by interaction with martian groundwater systems²¹. Trace-element and radiogenic-isotope patterns in Shergotty have been interpreted to reflect contamination by the martian crust^{28,29}, and the high oxidation state of Shergotty relative to some other shergottites³⁰ suggests that the assimilate may have been at least partly fluid. In any case, the inference that outgassing occurred in martian basaltic magma that erupted relatively recently (~175 Myr ago) argues that substantial water must have been delivered by magmas to the planet's surface in middle martian history. □

Methods

Shergotty USNM 321-2 was soaked for 2 h in a 1% mannitol solution, and rinsed in B-free distilled water to minimize surface contamination by B. The section was promptly coated with Au and analysed using the Oak Ridge National Laboratory Cameca ims-4f ion probe. The sample was bombarded with a 12.5-kV primary beam of $^{16}O^-$ ions, using a primary beam current of 10 nA and a beam size of ~20 μm . Positive secondary ions were extracted into the secondary mass spectrometer with a mass resolution of ~600 to eliminate the interference of Al^{13+} on 9Be . An energy offset of 80 eV was applied to eliminate molecular ions. Calibration curves were derived using multiple measurements of natural mineral and glass standards. Detection limits for Li, Be, B, Ce, Y and Ti were well below measured values.

The crystallization temperatures of pigeonite and augite in Shergotty were estimated using the pyroxenes in the experiment of longest duration. The 0.1-MPa experiment was run for 168 h and the 200-MPa experiment was run for 26 h under hydrous conditions.

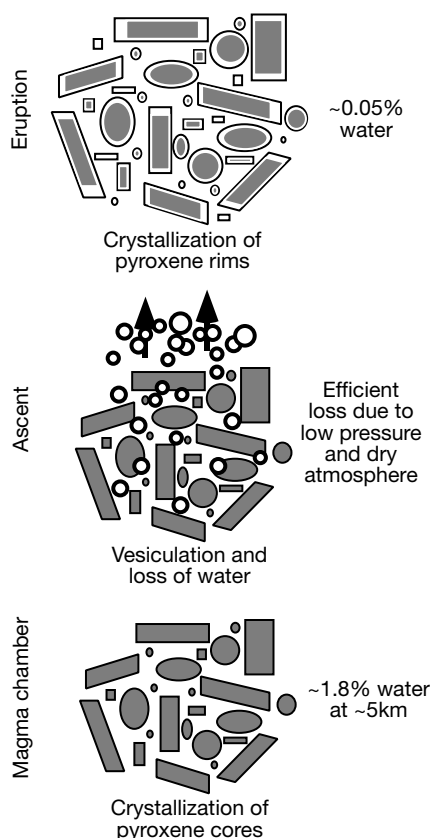


Figure 3 Schematic illustration of pyroxene crystallization and outgassing of the Shergotty parent magma at depth, during ascent, and on eruption.

These experimental durations produced relatively homogeneous crystalline products, as determined by electron microprobe. The 100-MPa experiment is at extreme conditions for hydrous experiments and could only be run for 9 h. The compositions of these pyroxenes varied significantly and were not used to estimate crystallization temperature.

Received 8 August; accepted 20 November 2000.

- Head, J. W. *et al.* Possible ancient oceans on Mars: Evidence from Mars Orbiter laser altimeter data. *Science* **286**, 2134–2137 (1999).
- Wänke, H. & Dreibus, G. Chemistry and accretion history of Mars. *Phil. Trans. R. Soc. Lond. A* **349**, 285–293 (1994).
- McSween, H. Y. Jr What we have learned about Mars from SNC meteorites. *Meteoritics* **29**, 757–779 (1994).
- Karlsson, H. R., Clayton, R. N., Gibson, E. K. Jr & Mayeda, T. K. Water in SNC meteorites: Evidence for a martian hydrosphere. *Science* **255**, 1409–1411 (1992).
- Stolper, E. M. & McSween, H. Y. Jr Petrology and origin of the shergottite meteorites. *Geochim. Cosmochim. Acta* **43**, 1475–1498 (1979).
- McSween, H. Y. Jr & Harvey, R. P. Outgassed water on Mars: Constraints from melt inclusions in SNC meteorites. *Science* **259**, 1890–1892 (1993).
- Watson, L. L., Hutcheon, I. D., Epstein, S. & Stolper, E. M. Water on Mars: Clues from deuterium/hydrogen and water contents of hydrous phases in SNC meteorites. *Science* **265**, 86–90 (1994).
- Mysen, B. O., Virgo, D., Popp, R. K. & Bertka, C. M. The role of H₂O in martian magmatic systems. *Am. Mineral.* **83**, 942–946 (1998).
- King, P. L., Hervig, R. L., Holloway, J. R., Vennemann, T. W. & Richter, K. Oxy-substitution and dehydrogenation in mantle-derived amphibole megacrysts. *Geochim. Cosmochim. Acta* **63**, 3635–3651 (1999).
- Seyfried, W. E. Jr, Janeky, D. R. & Mottl, M. J. Alteration of the oceanic crust: implications for geochemical cycles of lithium and boron. *Geochim. Cosmochim. Acta* **48**, 557–569 (1984).
- Brennan, J. M., Ryerson, F. J. & Shaw, H. F. The role of aqueous fluids in the slab-to-mantle transfer of boron, beryllium, and lithium during subduction: Experiments and models. *Geochim. Cosmochim. Acta* **62**, 3337–3347 (1998).
- Neal, C. R. & Taylor, L. A. A negative Ce anomaly in a peridotite xenolith: Evidence for crustal recycling into the mantle or mantle metasomatism? *Geochim. Cosmochim. Acta* **53**, 1035–1040 (1989).
- Hale, V. P. S., McSween, H. Y. Jr & McKay, G. Re-evaluation of intercumulus liquid composition and oxidation state for the Shergotty meteorite. *Geochim. Cosmochim. Acta* **63**, 1459–1470 (1999).
- McKay, G., Mikouchi, T., Le, L., Schwandt, C. & Hashimoto, M. The Shergotty paradox: An experimental perspective on intercumulus melt compositions. *Lunar Planet. Sci. [CD-ROM]* **31** (2000).
- McCoy, T. J. & Lofgren, G. E. Crystallization of the Zagami shergottite: An experimental study. *Earth Planet. Sci. Lett.* **173**, 397–411 (1999).
- Anderson, D. J., Lindsley, D. H. & Davidson, P. M. QUILF: A Pascal program to assess equilibria among Fe-Mg-Mn-Ti oxides, pyroxenes, olivine, and quartz. *Comput. Geosci.* **19**, 1333–1350 (1993).
- Hamilton, D. L., Burnham, C. W. & Osborn, E. F. The solubility of water and effects of oxygen fugacity and water content on crystallization in mafic magmas. *J. Petrol.* **5**, 21–39 (1964).
- Sisson, T. W. & Grove, T. L. Experimental investigations of the role of H₂O in calc-alkaline differentiation and subduction zone magmatism. *Contrib. Mineral. Petrol.* **113**, 143–166 (1993).
- Moore, G., Vennemann, T. & Carmichael, I. S. E. An empirical model for the solubility of H₂O in magmas to 3 kilobars. *Am. Mineral.* **83**, 36–42 (1998).
- Johnson, M. C., Rutherford, M. J. & Hess, P. C. Chassigny petrogenesis: Melt compositions, intensive parameters, and water contents of martian(?) magmas. *Geochim. Cosmochim. Acta* **55**, 349–366 (1991).
- Minitti, M. M. & Rutherford, M. J. Genesis of the Mars Pathfinder “sulfur-free” rock from SNC parental liquids. *Geochim. Cosmochim. Acta* **64**, 2535–2547 (2000).
- Wilson, L. & Head, J. W. Ascent and eruption of basaltic magma on the Earth and Moon. *J. Geophys. Res.* **86**, 2971–3001 (1981).
- Mouginis-Mark, P. J., Wilson, L. & Zuber, M. T. in *Mars* (eds Kieffer, H. H., Jakosky, B. M., Snyder, C. W. & Matthews, M. S.) 424–452 (Univ. Arizona Press, Tucson, 1992).
- Newman, S., Epstein, S. & Stolper, E. Water, carbon dioxide, and hydrogen isotopes in glasses from the ca. 1340 A.D. eruption of the Mono Craters, California: Constraints on degassing phenomena and initial volatile content. *J. Volcanol. Geotherm. Res.* **35**, 75–96 (1988).
- Mustard, J. F., Murchie, S., Erard, S. & Sunshine, J. M. In situ compositions of Martian volcanics: Implications for the mantle. *J. Geophys. Res.* **102**, 25605–25615 (1997).
- Bandfield, J. L., Hamilton, V. E. & Christensen, P. R. A global view of martian surface compositions from MGS-TES. *Science* **287**, 1626–1630 (2000).
- Lodders, K. & Fegley, B. Jr An oxygen isotope model for the composition of Mars. *Icarus* **126**, 373–394 (1997).
- Jones, J. H. Isotopic relationships among the shergottites, the nakhlites and Chassigny. *Proc. Lunar Planet. Sci. Conf.* **19**, 465–474 (1989).
- Longhi, J. Complex magmatic processes on Mars: Inferences from the SNC meteorites. *Proc. Lunar Planet. Sci. Conf.* **21**, 695–709 (1991).
- Herd, C. D. K. & Papke, J. J. Oxygen fugacity of the martian basalts from analysis of iron-titanium oxides: Implications for mantle-crust interaction on Mars. *Met. Planet. Sci.* **35**, A70 (2000).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

This work was partly supported by NASA.

Correspondence and requests for materials should be addressed to H.Y.McS. (e-mail: mcsween@utk.edu).

Observation of coherent optical information storage in an atomic medium using halted light pulses

Chien Liu^{*†}, Zachary Dutton^{*‡}, Cyrus H. Behroozi^{*†} & Lene Vestergaard Hau^{*†‡}

^{*}Rowland Institute for Science, 100 Edwin H. Land Boulevard, Cambridge, Massachusetts 02142, USA

[†]Division of Engineering and Applied Sciences, [‡]Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA

Electromagnetically induced transparency^{1–3} is a quantum interference effect that permits the propagation of light through an otherwise opaque atomic medium; a ‘coupling’ laser is used to create the interference necessary to allow the transmission of resonant pulses from a ‘probe’ laser. This technique has been used^{4–6} to slow and spatially compress light pulses by seven orders of magnitude, resulting in their complete localization and containment within an atomic cloud⁴. Here we use electromagnetically induced transparency to bring laser pulses to a complete stop in a magnetically trapped, cold cloud of sodium atoms. Within the spatially localized pulse region, the atoms are in a superposition state determined by the amplitudes and phases of the coupling and probe laser fields. Upon sudden turn-off of the coupling laser, the compressed probe pulse is effectively stopped; coherent information initially contained in the laser fields is ‘frozen’ in the atomic medium for up to 1 ms. The coupling laser is turned back on at a later time and the probe pulse is regenerated: the stored coherence is read out and transferred back into the radiation field. We present a theoretical model that reveals that the system is self-adjusting to minimize dissipative loss during the ‘read’ and ‘write’ operations. We anticipate applications of this phenomenon for quantum information processing.

With the coupling and probe lasers used in the experiment, the atoms are accurately modelled as three-level atoms interacting with the two laser fields (Fig. 1a). Under perfect electromagnetically-induced transparency (EIT) conditions (two-photon resonance), a stationary eigenstate exists for the system of a three-level atom and resonant laser fields, where the atom is in a ‘dark’, coherent superposition of states $|1\rangle$ and $|2\rangle$:

$$|D\rangle = \frac{\Omega_c|1\rangle - \Omega_p|2\rangle \exp[i(\mathbf{k}_p - \mathbf{k}_c) \cdot \mathbf{r} - i(\omega_p - \omega_c)t]}{\sqrt{\Omega_c^2 + \Omega_p^2}} \quad (1)$$

Here Ω_p and Ω_c are the Rabi frequencies, $\mathbf{k}_p$ and $\mathbf{k}_c$ the wavevectors, and ω_p and ω_c the optical angular frequencies of the probe and coupling lasers, respectively. The Rabi frequencies are defined as $\Omega_{p,c} \equiv e \mathbf{E}_{p,c} \cdot \mathbf{r}_{12,23} / \hbar$, where e is the electron charge, $\mathbf{E}_{p,c}$ are the slowly varying envelopes of probe and coupling field amplitudes, and $e \mathbf{r}_{12,23}$ are the electric dipole moments of the atomic transitions. The dark state does not couple to the radiatively decaying state $|3\rangle$, which eliminates absorption of the laser fields^{1–3}.

Atoms are prepared (magnetically trapped) in a particular internal quantum state $|1\rangle$ (Fig. 1a). The atom cloud is first illuminated by a coupling laser, resonant with the $|2\rangle$ – $|3\rangle$ transition. With only the coupling laser on and all atoms in $|1\rangle$, the system is in a dark state (equation (1) with $\Omega_p = 0$). A probe laser pulse, tuned to the $|1\rangle$ – $|3\rangle$ transition and co-propagating with the coupling laser, is subsequently sent through the atomic medium. Atoms within the pulse region are driven into the dark-state superposition of states $|1\rangle$ and $|2\rangle$, determined by the ratio of the instantaneous Rabi frequencies of

the laser fields (equation (1)).

The presence of the coupling laser field creates transparency, a very steep refractive index profile, and low group velocity, V_g , for the probe pulse^{1–10}. As the pulse enters the atomic medium, it is spatially compressed by a factor c/V_g whereas its peak electric amplitude remains constant during the slow-down^{4,7}.

The experiment is performed with the apparatus described in refs 4 and 11. Figure 1 shows the new optical set-up and atomic energy levels involved. A typical cloud of 11 million sodium atoms is cooled to 0.9 μK , which is just above the critical temperature for Bose–Einstein condensation. The cloud has a length of 339 μm in the z direction, a width of 55 μm in the transverse directions, and a peak density of 11 μm^{-3} . Those portions of the co-propagating probe and coupling laser beams that have passed through the 15- μm -diameter centre region of the cloud are selected and monitored simultaneously by two photomultiplier tubes (PMTs).

Figure 2a shows typical signals detected by the PMTs. The dashed curve is the measured intensity of the coupling laser, which is turned

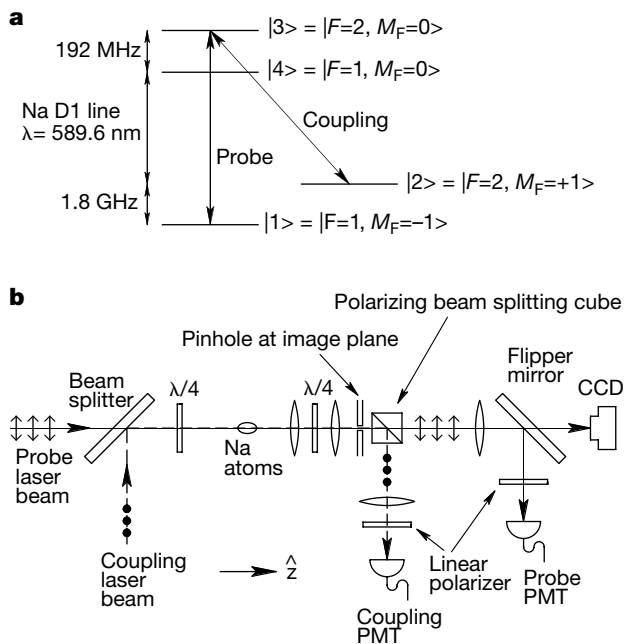


Figure 1 Experimental set-up and procedure. **a**, States $|1\rangle$, $|2\rangle$ and $|3\rangle$ form the three-level EIT system. The cooled atoms are initially magnetically trapped in state $|1\rangle = |3S, F=1, M_F=-1\rangle$. Stimulated photon exchanges between the probe and coupling laser fields create a 'dark' superposition of states $|1\rangle$ and $|2\rangle$, which renders the medium transparent for the resonant probe pulses. **b**, We apply a 2.2-mm diameter, σ^- -polarized coupling laser, resonant with the $|3S, F=2, M_F=+1\rangle \rightarrow |3P, F=2, M_F=0\rangle$ transition, and a co-propagating, 1.2-mm diameter σ^+ -polarized probe pulse tuned to the $|3S, F=1, M_F=-1\rangle \rightarrow |3P, F=2, M_F=0\rangle$ transition. The two laser beams start out with orthogonal linear polarizations (two-headed arrows and filled circles show the directions of linear polarization of the probe and coupling lasers, respectively). They are combined with a beam splitter, circularly polarized with a quarter-wave plate ($\lambda/4$), and then injected into the atom cloud. After leaving the cloud, the laser beams pass a second quarter-wave plate and regain their original linear polarizations before being separated with a polarizing beam-splitting cube. The atom cloud is imaged first onto an external image plane and then onto a CCD (charge-coupled device) camera. A pinhole is placed in the external image plane and positioned at the centre of the cloud image. With the pinhole and flipper mirror in place, only those portions of the probe and coupling laser beams that have passed through the central region of the cloud are selected and monitored simultaneously by two photomultiplier tubes (PMTs). States $|1\rangle$ and $|2\rangle$ have identical first-order Zeeman shifts so the two-photon resonance is maintained across the trapped atom clouds. Cold atoms and co-propagating lasers eliminate Doppler effects. However, off-resonance transitions to state $|4\rangle$ prevent perfect transmission of the light pulses in this case.

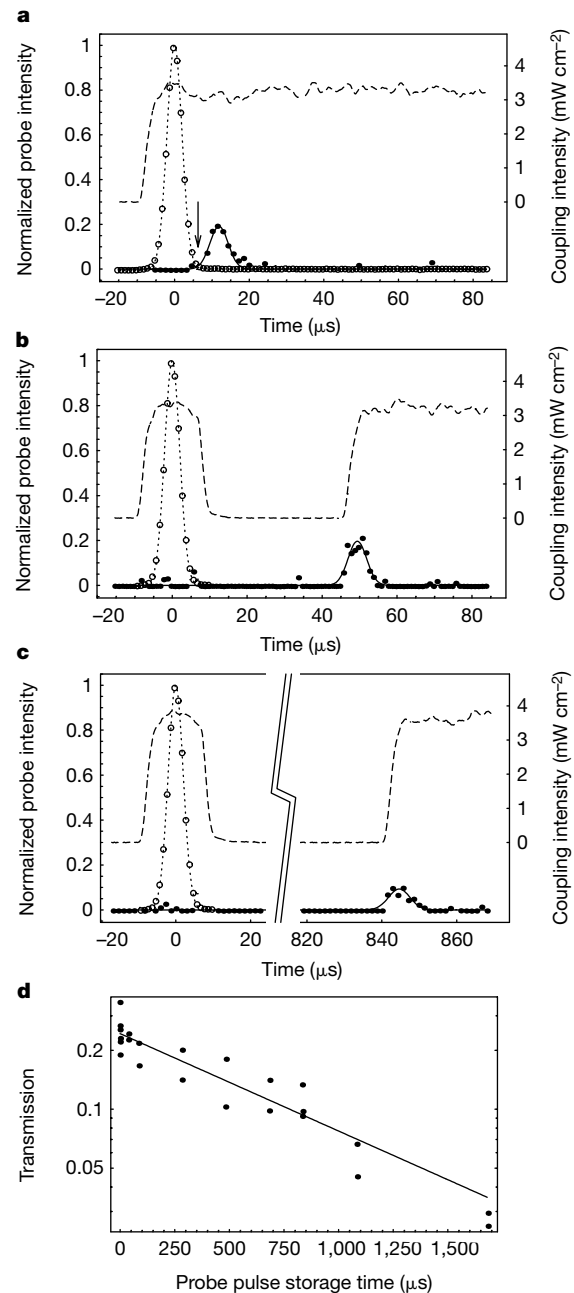


Figure 2 Measurements of delayed and revived probe pulses. Open circles (fitted to the dotted gaussian curves) show reference pulses obtained as the average of 100 probe pulses recorded in the absence of atoms. Dashed curves and filled circles (fitted to the solid gaussian curves) show simultaneously measured intensities of coupling and probe pulses that have propagated under EIT conditions through a 339- μm -long atom cloud cooled to 0.9 μK . The measured probe intensities are normalized to the peak intensity of the reference pulses (typically, $\Omega_p/\Omega_c = 0.3$ at the peak). **a**, Probe pulse delayed by 11.8 μs . The arrow at 6.3 μs indicates the time when the probe pulse is spatially compressed and contained completely within the atomic cloud. (The intersection of the back edge of the reference pulse and the front edge of the delayed pulse defines a moment when the tail of the probe pulse has just entered the cloud and the leading edge is just about to exit.) **b**, **c**, Revival of a probe pulse after the coupling field is turned off at $t = 6.3 \mu\text{s}$ and turned back on at $t = 44.3 \mu\text{s}$ and $t = 839.3 \mu\text{s}$, respectively. During the time interval when the coupling laser is off, coherent information imprinted by the probe pulse, is stored in the atomic medium. Upon subsequent turn-on of the coupling field, the probe pulse is regenerated through coherent stimulation. The time constants for the probe and coupling PMT amplifiers are 0.3 μs and 3 μs , respectively. The actual turn on/off time for the coupling field is 1 μs , as measured with a fast photodiode. **d**, Measured transmission of the probe pulse energy versus storage time. The solid line is a fit to the data, which gives a 1/e decay time of 0.9 ms for the atomic coherence.

on a few microseconds before the probe pulse. The open circles indicate a gaussian-shaped reference probe pulse recorded in the absence of atoms ($1/e$ full width is $5.70\ \mu\text{s}$). The filled circles show a probe pulse measured after it has passed through a cold atom cloud, and the solid curve is a gaussian fit to the data. The delay of this probe pulse, relative to the reference pulse, is $11.8\ \mu\text{s}$ corresponding to a group velocity of $28\ \text{m s}^{-1}$, a reduction by a factor of 10^7 from its vacuum value. The measured delay agrees with the theoretical prediction of $12.2\ \mu\text{s}$ based on a measured coupling Rabi frequency Ω_c of $2.57\ \text{MHz} \times 2\pi$ and an observed atomic column density of $3,670\ \mu\text{m}^{-2}$.

At time $t = 6.3\ \mu\text{s}$, indicated by the arrow in Fig. 2a, the probe pulse is spatially compressed and contained completely within the atomic cloud. The probe pulse in free space is $3.4\ \text{km}$ long and contains 27,000 photons within a $15\text{-}\mu\text{m}$ diameter at its centre. It is compressed in the atomic medium to match the size of the cloud ($339\ \mu\text{m}$), and the remaining optical energy in the probe field is only $1/400$ of a free-space photon. Essentially all of the probe energy has been transferred through stimulated emission into the coupling laser field and the atomic medium, and coherent optical information has been imprinted on the atoms (equation (1)).

To store this coherent information, we turn off the coupling field abruptly when the probe pulse is contained within the cloud. The stored information is read out at a later time by turning the coupling laser back on. A result is shown in Fig. 2b. The dashed curve shows the coupling laser's turn-off at $t = 6.3\ \mu\text{s}$ and its subsequent turn-on at $44.3\ \mu\text{s}$. The filled circles represent the measured probe intensity. As seen from the data, when the coupling laser is turned back on the probe pulse is regenerated: we can stop and controllably regenerate the probe pulse. Similar effects have been predicted in a

recent theoretical paper¹².

When the probe pulse is contained within the medium, the coherence of the laser fields is already imprinted on the atoms. As the coupling laser is turned off, the probe field is depleted to maintain the dark state (equation (1)) and (negligible) atomic amplitude is transferred from state $|1\rangle$ to state $|2\rangle$ through stimulated photon exchange between the two light fields. Because of the extremely low energy remaining in the compressed probe pulse, as noted above, it is completely depleted before the atomic population amplitudes have changed by an appreciable amount. When the coupling laser is turned back on, the process reverses and the probe pulse is regenerated through stimulated emission into the probe field. It propagates subsequently under EIT conditions as if the coupling beam had never been turned off.

During the storage time, information about the amplitude of the probe field is contained in the population amplitudes defining the atomic dark states. Information about the mode vector of the probe field is contained in the relative phase between different atoms in the macroscopic sample. The use of cold atoms minimizes thermal motion and the associated smearing of the relative phase during the storage time. (We obtain storage times that are up to 50 times larger than the time it takes an atom to travel one laser wavelength. As seen from equation (1), the difference between the wavevectors of the two laser fields determines the wavelength of the periodic phase pattern imprinted on the medium, which is 10^5 times larger than the individual laser wavelengths).

The regenerated probe pulse in Fig. 2b has the same shape as the 'normal' EIT pulse shown in Fig. 2a. Figure 2c shows a case where the optical coherence is stored in the atomic medium for more than $800\ \mu\text{s}$ before it is read out by the coupling laser. Here the amplitude

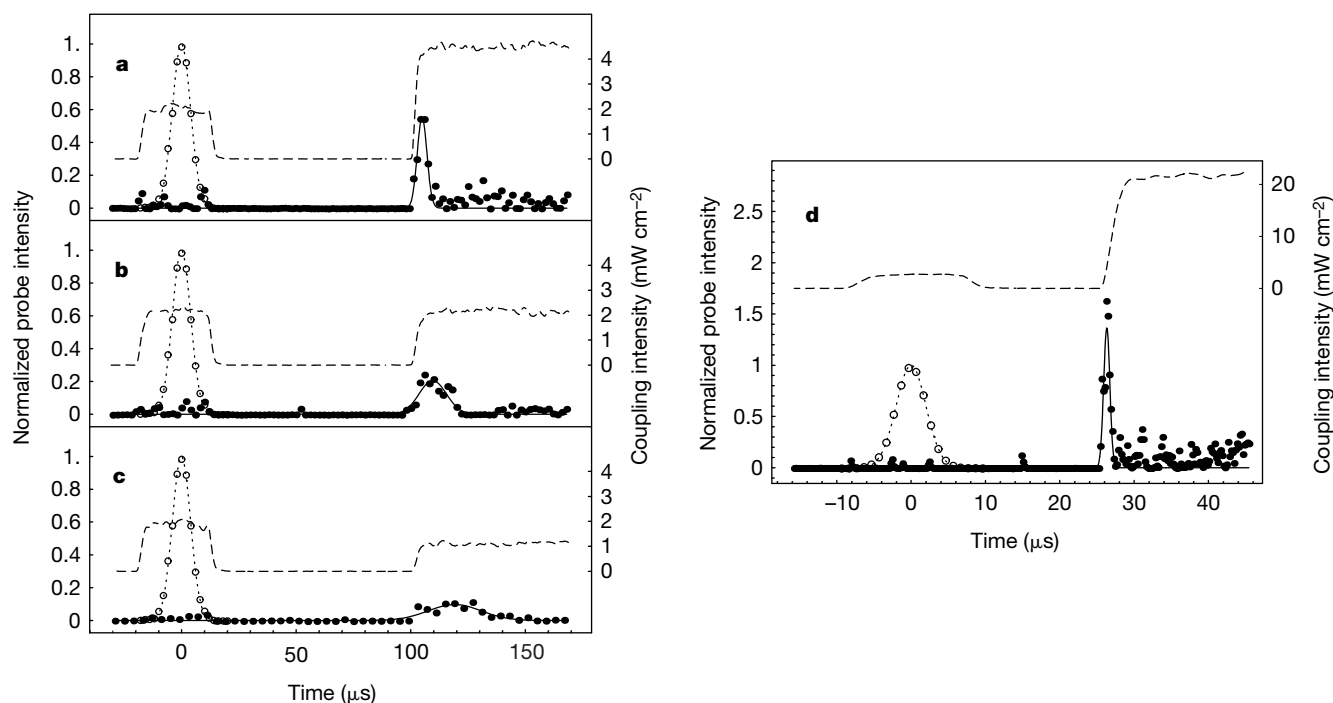


Figure 3 Measurements of revived probe pulses for varying intensities (I_{c2}) of the second coupling pulse. The intensity (I_{c1}) of the first coupling pulse is held constant. **a–c**, The figures are recorded for I_{c2}/I_{c1} ratios of 2., 1, and 0.5, respectively. A series of data show that the height and the inverse temporal width of the revived pulses are each proportional to I_{c2} . These observations are consistent with our physical picture. Because the atomic coherence dictates the ratio of the Rabi frequencies for the coupling and revived probe fields (equation (1)), the intensity of the regenerated probe pulse is proportional to the intensity of the coupling laser when it is turned back on. Furthermore, the spatial width of the revived pulse is determined by the distribution of the atomic coherence and is thus the

same as the spatial extent of the original compressed pulse. The group velocity of the probe pulse under EIT conditions is proportional to the coupling intensity^{4,7}. With a larger (smaller) I_{c2} , the revived probe pulse acquires a proportionally larger (smaller) group velocity, which causes its temporal width to be inversely proportional to I_{c2} . Panel **d** shows that the intensity of the revived probe pulse can exceed that of the original input pulse, in this instance by 40%. (The observed peak-to-peak fluctuation of laser intensity is less than 10%.) The energy in the revived probe pulses is the same in all panels **a–d**, owing to the fact that the total stored amplitude of state $|2\rangle$ atoms (available to stimulate photons into the probe field) is the same in all cases. Meanings of lines and symbols as in Fig. 2.

of the revived probe pulse is reduced compared to that of the pulse in Fig. 2b. Figure 2d shows the measured transmission for a series of pulses as a function of their storage time in the atom cloud. The data are consistent with an exponential decay with a $1/e$ decay time of 0.9 ms, comparable to the calculated mean free time of 0.5 ms between elastic collisions in the atom cloud with a density of $11 \mu\text{m}^{-3}$. Further studies of the decoherence mechanisms are planned but are beyond the scope of this Letter.

We have verified experimentally that the probe pulse is regenerated through stimulated rather than spontaneous emission. To do this, we prepared all atoms in state $|2\rangle$ and subsequently turned on the coupling laser alone. The coupling laser was completely absorbed for tens of microseconds without generating any signal in the probe PMT.

In Fig. 3a–c, we show three PMT signal traces recorded under similar conditions except that we vary the intensity, I_{c2} , of the coupling laser when it is turned back on. When I_{c2} is larger than the original coupling intensity, I_{c1} , the amplitude of the revived probe pulse increases and its temporal width decreases (Fig. 3a). For $I_{c2} < I_{c1}$, the opposite occurs (Fig. 3c). These results support our physical picture of the process. The stored atomic coherence dictates the ratio of the Rabi frequencies of the coupling and revived probe fields, as well as the spatial width of the regenerated pulse. In Fig. 3d we show that with a large I_{c2} , the peak intensity of the revived probe pulse exceeds that of the original input pulse by 40%.

Dissipationless pulse storage and revival processes are only possible if the ratio between the rates of dissipative and coherence-preserving events is small. When the coupling field is increased or decreased quickly compared to the duration of the probe pulse (τ) but slowly compared to $1/\Gamma$, this ratio is equal to (Z.D. and

L.V.H., manuscript in preparation)

$$\frac{2\Gamma}{\Omega_c^2 + \Omega_p^2} \left(\frac{\dot{\Omega}_p}{\Omega_p} - \frac{\dot{\Omega}_c}{\Omega_c} \right) \quad (2)$$

where Γ is the spontaneous decay rate from state $|3\rangle$. Our numerical simulations show that the probe field is constantly adjusting to match the changes in the coupling field in such a way that the terms in brackets in equation (2) nearly cancel^{13,14}. Even for turn-off times faster than $1/\Gamma$, we can show that there is no decay of the coherence between states $|1\rangle$ and $|2\rangle$ as long as $\tau \gg \Gamma/(\Omega_{c0}^2 + \Omega_{p0}^2)$; here Ω_{c0} and Ω_{p0} are the Rabi frequencies before the coupling turn-off. (The adiabatic requirement introduced in ref. 12 as necessary for non-dissipative behaviour is much too strict. That requirement would inevitably break down for low coupling laser powers during turn-on or turn-off.)

We have demonstrated experimentally that coherent optical information can be stored in an atomic medium and subsequently read out by using the effect of EIT in a magnetically trapped, cooled atom cloud. We have experimentally verified that the storage and read-out processes are controlled by stimulated photon transfers between two laser fields. Multiple read-outs can be achieved using a series of short coupling laser pulses. In Fig. 4a and b we show measurements of double and triple read-outs spaced by up to hundreds of microseconds. Each of the regenerated probe pulses contains part of the contents of the 'atomic memory', and for the parameters chosen, the memory is depleted after the second pulse and after the third pulse.

We believe that this system could be used for quantum information transfer; for example, to inter-convert stationary and flying qubits¹⁵. By injection of multiple probe pulses into a Bose–Einstein condensate—where we expect that most atomic collisions are coherence-preserving—and with use of controlled atom–atom interactions, quantum information processing may be possible during the storage time. □

Received 13 October; accepted 17 November 2000.

- Harris, S. E. Electromagnetically induced transparency. *Phys. Today* **50**, 36–42 (1997).
- Scully, M. O. & Zubairy, M. S. *Quantum Optics* (Cambridge Univ. Press, Cambridge, 1997).
- Arimondo, E. in *Progress in Optics* (ed. Wolf, E.) 257–354 (Elsevier Science, Amsterdam, 1996).
- Hau, L. V., Harris, S. E., Dutton, Z. & Behroozi, C. H. Light speed reduction to 17 metres per second in an ultracold atomic gas. *Nature* **397**, 594–598 (1999).
- Kash, M. M. *et al.* Ultraslow group velocity and enhanced nonlinear optical effects in a coherently driven hot atomic gas. *Phys. Rev. Lett.* **82**, 5229–5232 (1999).
- Budker, D., Kimball, D. F., Rochester, S. M. & Yashchuk, V. V. Nonlinear magneto-optics and reduced group velocity of light in atomic vapor with slow ground state relaxation. *Phys. Rev. Lett.* **83**, 1767–1770 (1999).
- Harris, S. E., Field, J. E. & Kasapi, A. Dispersive properties of electromagnetically induced transparency. *Phys. Rev. A* **46**, R29–R32 (1992).
- Grobe, R., Hioe, F. T. & Eberly, J. H. Formation of shape-preserving pulses in a nonlinear adiabatically integrable system. *Phys. Rev. Lett.* **73**, 3183–3186 (1994).
- Xiao, M., Li, Y.-Q., Jin, S.-Z. & Gea-Banacloche, J. Measurement of dispersive properties of electromagnetically induced transparency in rubidium atoms. *Phys. Rev. Lett.* **74**, 666–669 (1995).
- Kasapi, A., Jain, M., Yin, G. Y. & Harris, S. E. Electromagnetically induced transparency: propagation dynamics. *Phys. Rev. Lett.* **74**, 2447–2450 (1995).
- Hau, L. V. *et al.* Near-resonant spatial images of confined Bose-Einstein condensates in a 4-Dee magnetic bottle. *Phys. Rev. A* **58**, R54–R57 (1998).
- Fleischhauer, M. & Lukin, M. D. Dark-state polaritons in electromagnetically induced transparency. *Phys. Rev. Lett.* **84**, 5094–5097 (2000).
- Harris, S. E. Normal modes for electromagnetically induced transparency. *Phys. Rev. Lett.* **72**, 52–55 (1994).
- Fleischhauer, M. & Manak, A. S. Propagation of laser pulses and coherent population transfer in dissipative three-level systems: An adiabatic dressed-state picture. *Phys. Rev. A* **54**, 794–803 (1996).
- DiVincenzo, D. P. The physical implementation of quantum computation. Preprint quant-ph/0002077 at (<http://xxx.lanl.gov>) (2000).

Acknowledgements

We thank J. Golovchenko for discussions during which the idea of the rapid turn off and on of the coupling laser first emerged. We also thank M. Burns for critical reading of the manuscript. This work was supported by the Rowland Institute for Science, the Defense Advanced Research Projects Agency, the US Airforce Office of Scientific Research, and the US Army Research Office OSD Multidisciplinary University Research Initiative Program.

Correspondence and requests for materials should be addressed to C.L. (e-mail: chien@deas.harvard.edu).

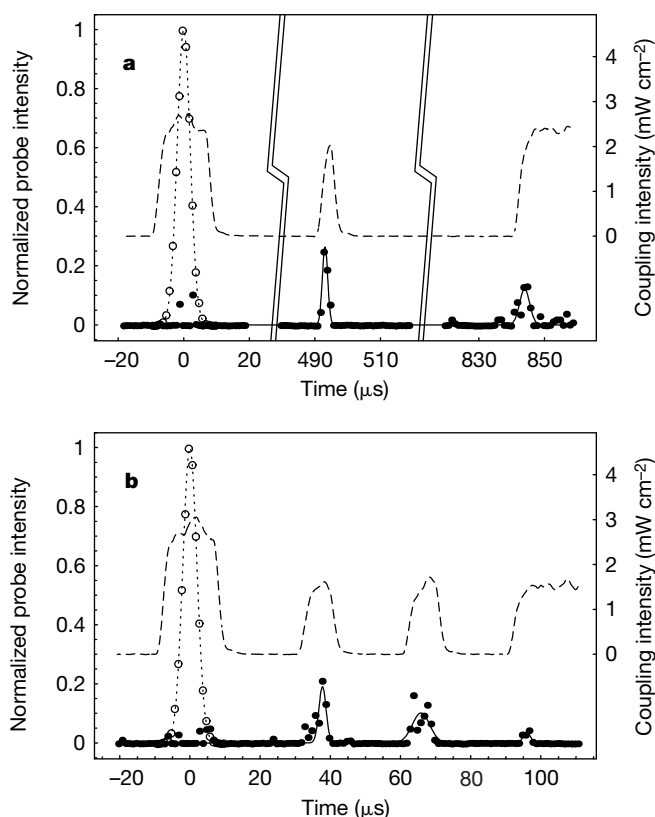


Figure 4 Measurements of double and triple read-out of the atomic memory. To deplete the atomic memory in these cases, we use two (a) and three (b) short coupling pulses. The total energy in the two (three) revived probe pulses is measured to be the same as the energy in the single revived probe pulse obtained with a single, long coupling laser pulse (as used in Figs 2 and 3). Meanings of lines and symbols as in Fig. 2.

Formation cross-sections of singlet and triplet excitons in π -conjugated polymers

M. Wohlgenannt^{†*}, Kunj Tandon^{†‡}, S. Mazumdar[§], S. Ramasesha[‡] & Z. V. Vardeny^{*}

^{*} Department of Physics, University of Utah, Salt Lake City, Utah 84112, USA

[‡] Solid State and Structural Chemistry Unit, Indian Institute of Science, Bangalore 560012, India

[§] Department of Physics, University of Arizona, Tucson, Arizona 85721, USA

[†] These authors contributed equally to this work

Electroluminescence in organic light-emitting diodes arises from a charge-transfer reaction between the injected positive and negative charges by which they combine to form singlet excitons that subsequently decay radiatively. The quantum yield of this process (the number of photons generated per electron or hole injected) is often thought¹ to have a statistical upper limit of 25 per cent. This is based on the assumption that the formation cross-section of singlet excitons, σ_s , is approximately the same as that of any one of the three equivalent non-radiative triplet exciton states, σ_t ; that is, $\sigma_s/\sigma_t \approx 1$. However, recent experimental² and theoretical³ work suggests that σ_s/σ_t may be greater than 1. Here we report direct measurements of σ_s/σ_t for a large number of π -conjugated polymers and oligomers. We have found that there exists a strong systematic, but not monotonic, dependence of σ_s/σ_t on the optical gap of the organic materials. We present a detailed physical picture of the charge-transfer reaction for correlated π -electrons, and quantify this process using exact valence bond calculations. The calculated σ_s/σ_t reproduces the experimentally observed trend. The calculations also show that the strong dependence of σ_s/σ_t on the optical gap is a signature of the discrete excitonic energy spectrum, in which higher energy excitonic levels participate in the charge recombination process.

The charge transfer (CT) reaction between the injected spin-half positively and negatively charged polarons (P^+ and P^-) proceeds by an intermediate metastable encounter complex (EC). The EC is a superposition of the overall eigenstate $|I\rangle$ of the initial reactant species and the overall eigenstate $|F\rangle$ of the final products of the CT reaction. Following the formation of an EC, conformational changes along reaction coordinates occur, leading to the formation of $|F\rangle$. For large yields $|EC\rangle$ must have simultaneously large overlaps with both $|I\rangle$ and $|F\rangle$. Here $|I\rangle = |P^+P^- \rangle$, $|F\rangle = |[S/T]\rangle + |G\rangle$, with $|P^\pm\rangle$ being the polaron eigenstates, $|[S/T]\rangle$ is either a neutral singlet or a neutral triplet excited state of one participant, and $|G\rangle$ is the ground state of the other participant. In the absence of electron correlation $|P^\pm\rangle$, $|[S/T]\rangle$ and $|G\rangle$ are all described by single configurations⁴. In particular, as the occupancies of the one-electron levels are the same for the singlet and triplet, $\sigma_s \approx \sigma_t$ for weak electron correlation.

The above description breaks down for intermediate electron correlations that are valid for π -conjugated polymers. First, single-configuration descriptions are no longer valid. The $|P^\pm\rangle$ are now superpositions of multiple configurations involving low (high) energy occupied (unoccupied) one-electron levels. Thus partial CT leads to multiple excited states of the EC supermolecule, each of which can decay to give the neutral states of the two components. Second, the correlated singlet exciton is necessarily ionic, whereas the correlated triplet exciton has a large covalent character⁵; there is also a substantial energy splitting between them^{6–8}. For large exciton production, the sum of the overlaps of the different eigenstates $|EC_j\rangle$ with $|I\rangle$ and $|F\rangle$ must be large. Because the initial polaronic species P^+ and P^- are necessarily ionic, the dominant $|EC_j\rangle$ must also be

ionic, and hence the most likely outcome of the CT reaction is the creation of product species, at least one of which is ionic. We conclude that $\sigma_s > \sigma_t$. Furthermore, the relative contributions of electron correlation and topology to the overall optical gap, which determine the ionicities of the ground and excited states, are strongly material-dependent; hence we expect σ_s/σ_t to be material-dependent from general considerations alone.

For systems which are light-emitting the quantum efficiency for electroluminescence (EL) is $\eta_{EL} = \eta_1\eta_2\eta_3$, where η_1 is the singlet emission quantum efficiency, η_2 is the fraction of the total number of excitons that are singlets, and η_3 is the probability that the injected electrons and holes find each other to form electron–hole pairs³. As both η_1 and $\eta_3 < 1$, it follows that $\eta_{EL} < \eta_2 = \eta_{max}$. Although $\sigma_s > \sigma_t$ by itself does not change the magnitude of η_2 from 1/4, in the presence of competing processes, faster rates will yield higher yields. One such competing process is the spin-lattice relaxation, which at room temperature is sufficiently fast that $\eta_{max} = \sigma_s/(\sigma_s + 3\sigma_t)$ (refs 3 and 9), rather than simply the statistical probability of obtaining one singlet exciton per three triplet excitons (see Supplementary Information). Thus the study of σ_s/σ_t in organic materials also provides information about η_{max} in organic light-emitting diodes (OLEDs).

We use continuous wave (CW) photoinduced absorption (PA) and photoinduced absorption detected magnetic resonance (PADMR) techniques (see Methods and ref. 10) to measure σ_s/σ_t in various π -conjugated polymer and oligomer thin films. For illustration, we choose the case of methylated laddertype poly-(*para*-phenylene) (mLPPP; see Fig. 1a, inset for its chemical structure¹¹). In Fig. 1a we show the CW PA spectrum of mLPPP at 80 K. The PA bands labelled P_1 and P_2 are the spectral signatures

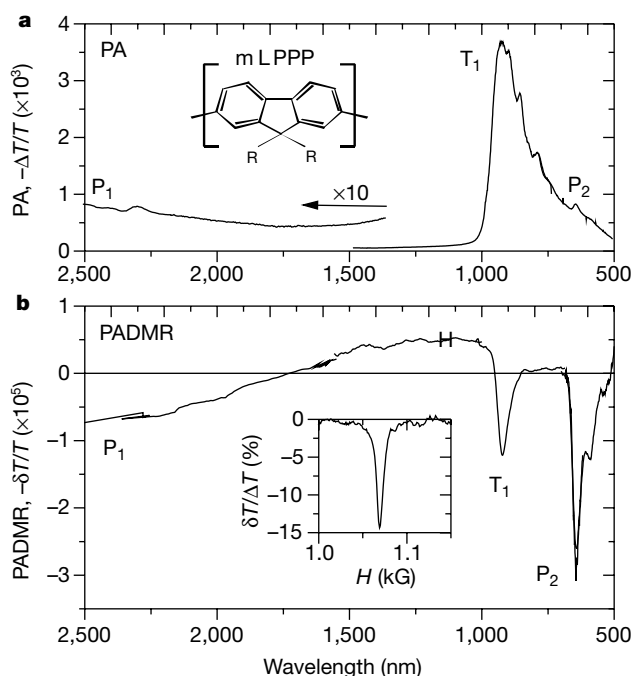


Figure 1 Spin-dependent recombination spectroscopy in π -conjugated polymers. For illustration the efficient blue emitter mLPPP (backbone is given in the inset to **a**) has been chosen. Its sharp spectral features make it ideal for spectroscopic studies¹⁰. **a**, The photoinduced absorption (PA) spectrum; **b**, the PA-detected magnetic resonance (PADMR) spectrum at constant magnetic field $H_{1/2} = 1.06$ kG. Both spectra **a** and **b** show two bands (P_1 and P_2) due to polarons, and one band (T_1) due to triplet absorption. The PA was measured at 80 K, the excitation was at the 457 nm line (~ 500 mW) of an Ar^+ laser; the PADMR spectrum was measured at 20 K. The inset in **b** shows the spin-half PADMR magnetic resonance at $H_{1/2} = 1.06$ kG detected at $\lambda = 2 \mu m$.

of polarons, whereas the PA band labelled T_1 corresponds to the triplet exciton¹². Under CW illumination conditions CT reactions occur between neighbouring P^+ and P^- . The CT reaction rate R_p between spin parallel pairs (both $\uparrow\uparrow$ and $\downarrow\downarrow$), is proportional to $2\sigma_T$, whereas the reaction rate R_{AP} between spin antiparallel pairs (both $\uparrow\downarrow$ and $\downarrow\uparrow$), is proportional to $(\sigma_S + \sigma_T)$, where the proportionality constant is the same in both cases. $\sigma_S > \sigma_T$, so $R_{AP} > R_p$, and spin polarization of the recombining polaron pairs is built up over time, such that spin parallel pairs prevail under steady state illumination conditions¹⁰.

Under saturated magnetic resonance conditions the Zeeman levels become equally populated, so that the pair densities with parallel and antiparallel spins are equal. Thus the effect of saturated magnetic resonance conditions is to increase the relative concentration of pairs with antiparallel spins whose reaction cross-section is larger, and as a consequence there is an overall decrease, δN , in the polaron population, N (see Fig. 1b inset). Figure 1b shows the wavelength-dependent PADMR spectrum of mLPPP where, in addition to a decrease in the polaron population (P_1 and P_2), a decrease in the triplet exciton population (T_1) is also measured. The negative triplet PADMR band confirms our statement that spin-half magnetic resonance decreases the population of neighbouring polaron pairs, especially those with parallel spins. The quantity $\delta T/\Delta T$ (see Methods) is then a direct measure of the fractional change in the overall photogenerated polaron population, $\delta N/N$.

The quantitative expression for $\delta N/N$ for distant pair kinetics under saturation conditions is given by¹⁰:

$$\delta N/N = -(R_p - R_{AP})^2/(R_p + R_{AP})^2 \quad (1)$$

From equation (1) and the proportionality relations above between R_p , R_{AP} and σ_S , σ_T , we obtain:

$$\frac{\sigma_S}{\sigma_T} = \frac{1 + 3|\delta T/\Delta T|^{1/2}}{1 - |\delta T/\Delta T|^{1/2}} \quad (2)$$

Thus the combination of PA and spin-half PADMR spectroscopy gives a direct measure of σ_S/σ_T . From Fig. 1b we get $\delta T/\Delta T \approx 14\%$ in mLPPP, corresponding to $\sigma_S/\sigma_T \approx 3.4$ from equation (2).

We performed similar PA and PADMR measurements for a large variety of π -conjugated polymers and oligomers. Our results are summarized in Fig. 2, where we plot the experimentally determined σ_S/σ_T as a function of the optical gap, E_g (see Supplementary Information for the experimental $\delta T/\Delta T$ values for each polymer). Individual $\sigma_S/\sigma_T > 1$ in all cases, giving $\eta_{\max} > 25\%$; but there is a very large variation between the materials, which has never been seen before, to our knowledge. $\sigma_S/\sigma_T \approx 2.2$ in poly(phenylene-vinylene) (PPV), which corresponds to $\eta_{\max}$ of 42%, in excellent agreement with refs 2 and 9, works in which $\eta_{\max}$ was directly measured from OLED operation.

Although disorder, morphology and chain length distributions are all important, the systematic behaviour in Fig. 2 precludes any of these from being the dominant factor. The single biggest difference between the materials shown originates from the relative contribution of electron correlation and topology to the overall optical gap. Importantly, the lowest-energy excitations of all such systems can be mapped onto those of linear polyene chains with artificially large bond alternations¹³. We therefore theoretically examine the CT reaction between two charged polyene chains as a function of varying bond alternation.

Our goal is to calculate σ_S/σ_T for the CT reaction $|P^+\rangle + |P^-\rangle \rightarrow |[S/T]\rangle + |G\rangle$, for the case of antiparallel spins. We consider two parallel polyene chains separated by about 4 Å (Fig. 3 inset), described by the hamiltonian $H = H_1 + H_2 + H_{12}$, where H_k ($k = 1, 2$) describe the individual chains and H_{12} is the interaction between them. H_k is the single-chain Pariser–Parr–Pople hamiltonian^{14,15}, with the difference that the bond alternation parameter, δ , in the intrachain hopping integral⁵, $t_{||}(1 \pm \delta)$, is

considered as a continuous variable. The interchain interaction is written as

$$H_{12} = \frac{1}{2} \sum_{i,j'} V_{ij'}(n_i - 1)(n_{j'} - 1) - t_{\perp} \sum_{i\sigma} (c_{i\sigma}^\dagger c_{i'\sigma} + \text{h.c.}) \\ + \sum_{i,j',\sigma} [ii'ii'](n_i + n_{i'})(c_{i\sigma}^\dagger c_{i'\sigma} + \text{h.c.}) \quad (3)$$

where h.c. is hermitian conjugate. Here i (j') are carbon atoms on chain 1 (2), $c_{i\sigma}^\dagger$ creates a π -electron of spin σ on carbon atom i , n_i is the total number of electrons on carbon atom i , $V_{ij'}$ is the interchain Coulomb interaction, i, i' are the nearest interchain neighbours with $t_{\perp}$ the corresponding hopping integral (see Fig. 3 inset), and $[ii'ii']$ is the bond-charge repulsion¹⁶. The second and the third terms in equation (3) promote interchain CT. Here $t_{||}(t_{\perp})$ is the intrachain (interchain) or parallel (perpendicular) hopping integral.

Exact calculations (see Methods of relative singlet and triplet yields were performed for pairs of ethylene, butadiene and hexatriene. Calculations for such small systems are not expected to give a quantitatively valid picture, which is not possible anyway, given the uncertainties in the parameters of H_{12} . Nevertheless, we believe that it is far more important to consider the full effects of electron correlation, rather than perform approximate calculations for longer polyenes and oligomers of the experimental systems that do not capture the many-body nature of $|P^{\pm}\rangle$, $|S\rangle$ and $|T\rangle$. In Fig. 3 we show the results of our model calculation for two hexatriene chains, where the calculated σ_S/σ_T ratio has been plotted versus the bond alternation parameter δ , for the specific case of $t_{\perp} = [ii'ii'] = 0.1$ eV. The calculated σ_S/σ_T in Fig. 3 exhibits three distinct peaks, with the first (relatively small) occurring at $\delta = 0$. Qualitatively similar results were also obtained with butadiene (see Supplementary Information).

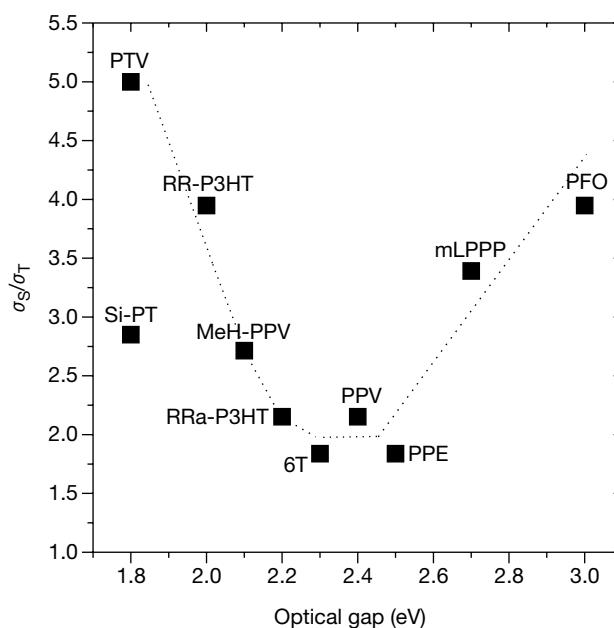


Figure 2 The experimentally determined σ_S/σ_T ratio for several π -conjugated polymers and oligomers as a function of the optical gap, E_g . σ_S/σ_T was evaluated using equation (2) from the measured $\delta T/\Delta T$ ratio for polarons in the PADMR and PA spectra, respectively. The dotted line is a guide to the eye. To ensure that δT were measured at saturated microwave absorption condition, we verified that the microwave field may be attenuated by 5 dB without reducing the observed $\delta T/\Delta T$ signals (see Supplementary Information). The values of $\delta T/\Delta T$ do not change with laser intensity or modulation frequency (see Supplementary Information) and are thus an intrinsic property of the individual polymer material.

The underlying mechanism for the occurrence of multiple peaks in Fig. 3 is as follows. For non-zero electron correlations the CT process need not lead to the lowest singlet or triplet exciton level. The correlated $|P^{\pm}\rangle$ wavefunctions, which consist of numerous configurations, change with increasing δ , and so both the individual $|\text{EC}_j\rangle$ and the specific eigenstate j that dominates the CT process evolve continuously. As a consequence of the changing character of $|P^{\pm}\rangle$, different $n^1\text{B}_u$ exciton states dominate the CT yields at different δ , where n is the exciton quantum number. Furthermore, because the spectrum of our finite system is discrete, there exist regions of δ where neither of the two consecutive $n^1\text{B}_u$ and $(n+1)^1\text{B}_u$ excitons give large yields. The three peaks in Fig. 3 at $\delta = 0, 0.2$ and 0.6 spectroscopically correspond to the 1^1B_u , the 2^1B_u and the 3^1B_u states, respectively (see Supplementary Information).

Similar effects also occur in the polymers (Fig. 2), where topological one-electron contributions to E_g increase from polyfluorene (PFO) to poly(thienylene vinylene) (PTV). Two aspects of the polymers allow direct comparison to our short-chain calculations. First, enhanced contribution of the topological gap to E_g can be simulated by increasing δ within effective linear-chain models¹³. Second, long-range Coulomb interactions in the Pariser–Parr–Pople hamiltonian^{14,15} give several excitons even in the long-chain limit¹⁷. Thus the qualitative effect of increased topological contribution to the optical gap is the same as increasing δ in our model calculation in Fig. 3, that is, the dominant CT product for large topological gap is a $n^1\text{B}_u$ exciton state with n larger than 1. Experimentally, the initial 1^1B_u state that is the product of the CT reaction in mLPPP and PFO has a quantum number that is larger than the initial 1^1B_u state in the case of PTV and regio-regular poly(3-hexylthiophene) (RR-P3HT), with α -hexathiophene (6-HT), PPV and poly(phenylene-ethylene) (PPE) lying in a region of parameter space that corresponds to one of the valleys in Fig. 3. It is difficult to assign δ values to all systems in Fig. 2. Previous work has assigned $\delta \geq 0.2$ to PPV (ref. 13), and with moderate shifts of the peaks in Fig. 3 with chain length to lower δ (see Supplementary Information), the valley in σ_S/σ_T may well occur near $\delta \approx 0.2$, with the next peak occurring at $\delta \approx 0.5$. We therefore conclude that the obtained systematic variation of the σ_S/σ_T shown in Fig. 2 is a distinct signature of the excitonic energy spectrum of the lowest-lying excited states in π -conjugated polymers.

We have thus discovered, to the best of our knowledge, a new experimental technique to determine the σ_S/σ_T ratio in π -conjugated polymers and oligomers using a spin-dependent recombination spectroscopy. The technique is very general, and may be used in other organic or inorganic materials that have long-lived charge excitations, particularly in determining the maximum EL yield in

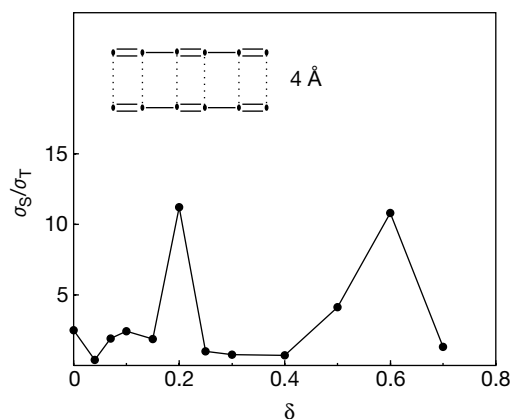


Figure 3 The calculated σ_S/σ_T for two parallel chains of hexatriene as a function of hypothetical bond alternation δ . The inset shows the actual arrangement of the chains chosen for the calculation.

OLEDs made from the individual systems. Theoretically, our approach takes into account the full effects of electron correlation. The curious variation obtained in σ_S/σ_T with E_g is a novel electron correlation effect, that gives a new perspective on the electron–hole recombination in organic semiconductors. □

Methods

The CW PA and spin-half PADMR techniques

The CW PA technique measures the excited state absorption spectrum of long-lived photoexcitations, such as triplet excitons and charged polarons. Two light beams are used; for the excitation beam we used an Ar⁺ laser, the intensity of which was modulated with a chopper. We measured the pump-beam-induced changes, ΔT , in the probe beam (tungsten halogen lamp) transmission, T , using a monochromator and a combination of solid state detectors. The concentration, N , of the long-lived species is proportional to the corresponding PA intensity ($-\Delta T/T$).

The PADMR technique measures the effect of spin-half magnetic resonance on the steady state population of photogenerated polarons and triplets. The experiment consists, in addition to the PA set-up, of a magnetic resonance part (microwave source and resonator, superconducting magnet). The sample is put inside the resonator and cooled by liquid helium. The PADMR set-up allows both PA and PADMR measurements under identical conditions. The change, δN , in the steady state population is proportional to the corresponding PADMR intensity ($-\delta T/T$, where δT is the resonant change in transmission). Experimentally the PADMR resonance is achieved by matching the energy splitting between the two Zeeman levels at magnetic field $H_{1/2}$ to the photon energy of an intense microwave field. For the 3-GHz microwave resonator used here, $H_{1/2}$ (corresponding to a g -value of 2) amounts to 1,060 G.

Computational technique

The approach for calculating σ_S/σ_T for the two charge chains is as follows. We use a time-dependent Schrödinger approach to calculate the time evolution of the initial state $\psi(0) = |P^+|P^- \rangle$, when operated by the overall hamiltonian H . The time evolution is done following a discretized procedure (see Supplementary Information), and after each evolution step, the evolved state $\psi(t)$ (which contains all the $|\text{EC}_j\rangle$) is projected onto the product of the eigenstates of the neutral systems. The yield for a given pair of product states is the overlap $|\langle \psi(t) | m, n \rangle|^2$, where $|m, n\rangle = |m\rangle \times |n\rangle$, with one of the two components $|m\rangle$ and $|n\rangle$ corresponding to $|\text{S}/\text{T}\rangle$ and the other to $|\text{G}\rangle$. The procedure was checked by performing the calculations first for the one-electron limit, where all yields are known.

Received 21 September; accepted 5 December 2000.

1. Friend, R. H. *et al.* Electroluminescence in conjugated polymers. *Nature* **397**, 121–128 (1999).
2. Cao, Y., Parker, I. D., Yu, G., Zhang, C. & Heeger, A. J. Improved quantum efficiency for electroluminescence in semiconducting polymers. *Nature* **397**, 414–417 (1999).
3. Shuai, Z., Beljonne, D., Silbey, R. J. & Bredas, J. L. Singlet and triplet exciton formation rates in conjugated polymer light-emitting diodes. *Phys. Rev. Lett.* **84**, 131–134 (2000).
4. Campbell, D. K., Bishop, A. R. & Fesser, K. Polarons in quasi-one-dimensional systems. *Phys. Rev. B* **26**, 6862–6874 (1982).
5. Soos, Z. G. & Ramasesha, S. Valence bond theory of linear Hubbard and Pariser–Parr–Pople models. *Phys. Rev. B* **29**, 5410–5422 (1984).
6. Lavrentiev, M. Yu., Barford, W., Martin, S. J., Daly, H. & Bursill, R. J. Theoretical investigation of the low-lying electronic structure of poly(p-phenylene vinylene). *Phys. Rev. B* **59**, 9987–9994 (1999).
7. Röhlfing, M. & Louie, S. G. Optical excitations in conjugated polymers. *Phys. Rev. Lett.* **82**, 1959–1962 (1999).
8. Monkman, A. P. *et al.* The triplet state of the ladder-type methyl-poly(p-phenylene) as seen by pulse radiolysis-energy transfer. *Chem. Phys. Lett.* **327**, 111–116 (2000).
9. Kim, J., Ho, P. K. H., Greenham, N. C. & Friend, R. H. Electroluminescence emission pattern of organic light-emitting diodes: Implications for device efficiency calculations. *J. Appl. Phys.* **88**, 1073–1081 (2000).
10. Vardeny, Z. V. & Wei, X. in *Handbook of Conducting Polymers II* (eds Skotheim, T. A., Elsenbaumer, R. L. & Reynolds, J. R.) 639–666 (Marcel Dekker, New York, 1997).
11. Graupner, W. *et al.* Direct observation of ultrafast field-induced charge generation in ladder-type poly(para-phenylene). *Phys. Rev. Lett.* **81**, 3259–3262 (1998).
12. Wohlgenannt, M., Graupner, W., Leising, G. & Vardeny, Z. V. Photogeneration and recombination processes of neutral and charged excitations in films of a ladder-type poly(para-phenylene). *Phys. Rev. B* **60**, 5321–5330 (1999).
13. Soos, Z. G., Ramasesha, S. & Galvão, D. S. Band to correlated crossover in alternating Hubbard and Pariser–Parr–Pople chains: Nature of the lowest singlet excitation of conjugated polymers. *Phys. Rev. Lett.* **71**, 1609–1612 (1993).
14. Pariser, R. & Parr, R. G. A semi-empirical theory of the electronic spectra and electronic structure of complex unsaturated molecules II. *J. Chem. Phys.* **21**, 767–776 (1953).
15. Pople, J. A. Electron interaction in unsaturated hydrocarbons. *Trans. Faraday Soc.* **49**, 1375–1385 (1953).
16. Campbell, D. K., Gammel, J. T. & Loh, E. Y. Bond-charge Coulomb repulsion in Peierls–Hubbard models. *Phys. Rev. B* **38**, 12043–12046 (1988).
17. Gallagher, F. & Mazumdar, S. Excitons and optical absorption in one-dimensional extended Hubbard models with short- and long-range interactions. *Phys. Rev. B* **56**, 15025–15039 (1997).

Supplementary information is available on Nature's World-Wide Web site or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank E.J.W. List for supplying the mLPPP polymer, and X.M. Jiang and E.J.W. List for their support in performing the experiments. The work at the University of Utah was supported in part by the DOE and the NSF. Work in Arizona was partially supported by the ONR through the MURI centre (CAMP) at the University of Arizona. Work in Bangalor was funded by CSIR, India.

Correspondence and requests for materials should be addressed to Z.V.V. (e-mail: val@physics.utah.edu).

Riverine export of aged terrestrial organic matter to the North Atlantic Ocean

Peter A. Raymond*† & James E. Bauer*

*School of Marine Science, College of William and Mary, Gloucester Point, Virginia 23062, USA

† Present address: Marine Biological Laboratory, Ecosystems Center, Woods Hole, Massachusetts 02543, USA

Global riverine discharge of organic matter represents a substantial source of terrestrial dissolved and particulate organic carbon to the oceans^{1,2}. This input from rivers is, by itself, more than large enough to account for the apparent steady-state replacement times of 4,00–6,000 yr for oceanic dissolved organic carbon^{3–5}. But paradoxically, terrestrial organic matter, derived from land plants, is not detected in seawater and sediments in quantities that correspond to its inputs^{6–8}. Here we present natural ¹⁴C and ¹³C data from four rivers that discharge to the western North Atlantic Ocean and find that these rivers are sources of old (¹⁴C-depleted) and young (¹⁴C-enriched) terrestrial dissolved organic carbon, and of predominantly old terrestrial particulate organic carbon. These findings contrast with limited earlier data⁹ that suggested terrestrial organic matter transported by rivers might be generally enriched in ¹⁴C from nuclear testing, and hence newly produced. We also find that much of the young dissolved organic carbon can be selectively degraded over the residence times of river and coastal waters, leaving an even older and more refractory component for oceanic export. Thus, pre-ageing and degradation may alter significantly the structure, distributions and quantities of terrestrial organic matter before its delivery to the oceans.

We sampled four rivers that discharge to the western North Atlantic Ocean: the Amazon (Brazil), Hudson (New York, USA), York (Virginia, USA) and Parker (Massachusetts, USA). The Amazon was sampled during base flow in 1991, and the three US rivers were sampled during moderate to high flow regimes from 1996 to 1998. The four sampled rivers range over four orders of magnitude in freshwater discharge (160,000, 375, 70 and 11 m³ s^{–1}, respectively) and watershed size (7,050,000, 21,000, 4,350 and 609 km², respectively). All water samples were taken from tidal freshwater sites except the Parker River sample, which was taken from above a dam feeding the Parker estuary. We report radio-carbon (¹⁴C) signatures for two operationally defined size fractions of organic matter: dissolved organic carbon (DOC; the fraction passing through a 0.7-μm glass-fibre filter) and particulate organic carbon (POC; the fraction collected on a 0.7-μm glass-fibre filter). Together these two size classes comprise the total organic carbon in these river waters. We also present for comparison data from four other western North Atlantic rivers¹⁰ collected previously.

Mean DOC and POC concentrations for all rivers were 357 ± 280 μM and 95 ± 33 μM, respectively (Tables 1 and 2). The ¹⁴C of DOC exported from these rivers was highly variable, ranging over 415‰ (¹⁴C is defined in Table 1.) The Hudson River DOC had the most depleted ¹⁴C value at –158‰ (equivalent radiocarbon age of 1,384 years before present, yr BP), while the DOC from the Susquehanna and Rappahannock rivers was also low in ¹⁴C-DOC, with corresponding ages of 680 and 766 yr BP (Table 1). In contrast, the ¹⁴C of DOC in the York, Parker, Potomac and Amazon rivers all contained ‘bomb’ ¹⁴C (that is, ¹⁴C > –50‰), introduced by nuclear weapons testing in the 1950s and 1960s.

The corresponding ¹⁴C values for POC (Table 2) were all significantly depleted compared to DOC (Table 1). The average ¹⁴C-POC values for the Amazon, Hudson, York and Parker rivers were –145, –437, –68 and –138‰ and had corresponding average radiocarbon ages of 1,258, 4,609, 690 and 1,210 yr BP, respectively; the greatest ages for POC from these four rivers were 1,258, 4,763, 1,690 and 1,696 yr BP, respectively (Table 2). This suggests that on average, the bulk POC discharged from these rivers was photo-synthetically fixed several millennia ago, and a significant fraction of this POC aged on land and in river basins for hundreds to thousands of years. Very old riverine POC (>10,000 yr BP) has also been found in a subtropical mountainous river in Taiwan¹¹. The average radiocarbon ages of POC in the present study were 1,260, 3,225, 690 and 1,210 years older than DOC in the Amazon, Hudson, York and Parker rivers, respectively (Tables 1 and 2). The ¹⁴C values of Amazon DOC and POC were ~200‰ lower than values found further upriver in 1984 for dissolved humic substances and POC⁹, and reflects a decrease in atmospheric ¹⁴C-CO₂ of 140‰ over the same period¹², spatial variation in POC sources¹³, as well as our

Table 1 Concentrations and isotope data for riverine DOC

River	Date	DOC (μM)	Δ ¹⁴ C (‰)	Radiocarbon age (yr BP)	δ ¹³ C (‰)
Amazon	11/91	235	28±6	Modern	–28.0
Hudson	06/98	196	–158±7	1,384	–25.5
York	09/96	701	216±5	Modern	–28.8
	11/96	443	208±5	Modern	–27.9
	03/97	390	257±7	Modern	–28.0
	06/97	435	159±5	Modern	–28.0
Parker	06/98	986	109±6	Modern	–28.3
Potomac*	1972	364	+161	Modern	–30.9
Susquehanna*	1972	292	–81	680	ND
Rappahannock*	1973	125	–91	766	–31.9
James*	1973	167	+42	Modern	–28.0

Values of Δ¹⁴C are expressed as the deviation in parts per thousand (‰) from the ¹⁴C activity of nineteenth century wood. δ¹³C values are expressed as ((R_{sample}/R_{standard}) – 1) × 10³ in ‰, where R = ¹³C/¹²C, and the standard is the Pee Dee Belemnite. DOC samples (100 ml of 0.7-μm-filtered river water) were oxidized to CO₂ by high-energy (2,400 V) ultraviolet (UV) irradiation for 2 h (refs 3, 5). The CO₂ samples were then converted to graphite and analysed for ¹⁴C by accelerator mass spectrometry (AMS)²⁰. All Δ¹⁴C values were corrected for sample δ¹³C (ref. 31). Errors (± 1σ) associated with Δ¹⁴C AMS analyses averaged ±6‰ (±60 years for radiocarbon age), while those for δ¹³C analyses averaged ±0.1‰. Concentrations of DOC were determined as part of the UV oxidation and CO₂ purification procedure, with quantification by a positive pressure (Baratron) gauge. The average error (± 1σ) for DOC concentrations determined by this method was ±1 μM. ND, not determined.

* Values from ref. 10, standard deviations not available.

Table 2 Concentrations and isotope data for riverine POC

River	Date	POC (μM)	$\Delta^{14}\text{C}$ (‰)	Radiocarbon age (yr BP)	$\delta^{13}\text{C}$ (‰)
Amazon	11/91	52	-145 ± 6	1,258	-25.6
Hudson	06/98	118	-447 ± 21	4,763	ND
	10/98	43	-426 ± 27	4,456	ND
York	09/96	70	24 ± 13	Modern	ND
	11/96	30	-38 ± 18	316	-28.2
	06/98	218	-190 ± 6	1,690	-30.0
Parker	06/98	208	-85 ± 20	715	-30.0
	10/98	75	-190 ± 25	1,696	-33.7

Suspended POC samples were collected by filtering river water through a baked ($>550^\circ\text{C}$) 0.7- μm glass-fibre filter. Filters were exposed to fuming HCl to remove carbonates before analysis, thoroughly dried, and processed by sealed quartz-tube combustion (900°C using a CuO/Ag metal catalyst) to produce CO_2 (ref. 32). Procedures and errors associated with POC, $\Delta^{14}\text{C}$ and $\delta^{13}\text{C}$ analyses averaged $\pm 17\%$ (± 160 yr for radiocarbon age), while those for ^{13}C analyses averaged $\pm 0.1\%$. ND, not determined.

analysis of the total DOC pool (as opposed to just the humic fraction).

The origins of natural riverine organic matter include terrestrial soils, continental and marine sedimentary rocks, and autochthonous production², although the dominant source of DOC and POC to most rivers is of terrestrial origin^{2,14}. For this study, the $\Delta^{14}\text{C}$ of dissolved inorganic carbon available for photosynthesis indicates that it was enriched by at least 100‰ compared to DOC and POC (P.A.R. and J.E.B., unpublished data). This enrichment and the low concentrations of chlorophyll *a* ($< 5.0 \mu\text{g l}^{-1}$) in most samples both argue against significant contributions from autochthonous phytoplankton production. Soil profiles of temperate forests indicate the presence of distinct organic carbon components that range from modern in the litterfall and upper-soil horizons to $> 9,000$ years in age in deeper horizons^{15–18}. In soils, the DOC of pore water is often enriched in ^{14}C with respect to the atmosphere¹⁸, and solubilized DOC is younger by hundreds to thousands of years than the residual organic matter (that is, POC)^{16,17}. Therefore, because soils are the main source of riverine organic matter^{2,14}, our data suggest that when there is low algal input, riverine POC contains a significant quantity of old, non-hydrolysable, recalcitrant soil organic matter originating from deeper soil horizons transported to rivers via erosion events. Conversely, riverine DOC consists of a somewhat younger, hydrolysable fraction that is transported to rivers following its continuous leaching from soils and ^{14}C -enriched litter fall. For all rivers, the $\delta^{13}\text{C}$ values for DOC and POC were depleted (-33.7‰ to -25.5‰ ; Tables 1 and 2), which is consistent with inputs of terrestrially derived plant matter and/or old (Mesozoic or older) sedimentary sources¹⁹.

In addition to the inherent differences in the ages of soluble versus insoluble (that is, particulate) organic matter in parent rocks and soils, the processes governing DOC and POC ages and variability in rivers include the weathering and denudation regimes of the watersheds, the absolute ages of parent rocks and soil organic matter, and the residence times of DOC and POC in feeder streams and rivers themselves. Some of the oldest POC was found in the Amazon and Hudson rivers (Table 2), which have relatively high-relief headwater regions^{20,21} where there are thin soils continuously eroded by mechanical weathering. The contribution of significant amounts of ^{14}C -depleted sedimentary carbon^{2,11}, present in both these watersheds^{20,21}, could explain the old ages of POC in such rivers. Alternatively, rivers draining areas of low relief with thicker, younger soils (for example, the York and Parker rivers) are dominated by chemical weathering (that is, solubilization of organic matter), and may have smaller relative contributions from deeper/older soil profiles and sedimentary rocks. We note that while rivers such as the Amazon have a significant contribution of suspended solids²² and POC (ref. 13) from the Andean highlands, they would also contain modern DOC and POC from lowland floodplains¹³.

Our finding of much older POC than DOC in all of these rivers is

also consistent with a reservoir or storage effect associated with POC 'spiralling' (that is, suspension/re-deposition) in streams and rivers²³. Because POC is believed to consist of soluble organic matter adsorbed on mineral surfaces^{24,25}, the storage effects for POC and mineral particles may be similar. Thus, geomorphic and hydrologic variations between watersheds will have an important effect on the absolute ages of riverine DOC and POC, and will contribute to the pattern of older POC in rivers. Another factor potentially contributing to the old ages of POC is human disturbance. In the Hudson, for instance, the chief source of POC is from agricultural fields²⁶, where older soil horizons have been tilled to the surface and are susceptible to erosion.

An important fate of DOC in rivers and estuaries is bacterial degradation^{27,28}. We conducted long-term, dark incubations (2–12 months) with York River water to investigate how bacterial utilization of DOC in rivers alters the $\Delta^{14}\text{C}$ signatures of riverine DOC prior to its discharge to the coastal ocean. The $\Delta^{14}\text{C}$ values of DOC decreased in all incubations (Fig. 1a), indicating that there is preferential removal of ^{14}C -enriched (that is, young) labile DOC by heterotrophic bacteria. In the one-year incubation, 63% of the initial bulk DOC pool ($\Delta^{14}\text{C} = 160\text{‰}$) was degraded, and the residual fraction had a $\Delta^{14}\text{C}$ of -65‰ (or ~ 540 yr in age). These findings indicate that riverine DOC consists of a larger pool of biologically labile, ^{14}C -enriched (that is, young) DOC, and a smaller pool of biologically refractory, ^{14}C -depleted (that is, old) DOC. The $\Delta^{14}\text{C}$ of the refractory fraction of riverine DOC is also directly related to the amount of DOC utilized by bacteria (Fig. 1b). Thus we predict that, as younger DOC is degraded, the remaining DOC (which is available for export to the coastal ocean and beyond) will be much older than the original material. When measuring total $\Delta^{14}\text{C}$ -DOC, this old pool may be further concealed because of the intrusion of bomb ^{14}C .

The radiocarbon signatures of DOC and POC, and the shift towards older residual DOC following bacterial degradation, suggest that organic matter exported from rivers to the oceans is older, and much more variably aged, than previously thought. The limited previous data on ages of riverine organic matter⁹ have led to the presumption that rivers in general transport—and export to the oceans—young DOC and POC originating from terrestrial sources. Approximately 40% of the organic matter transported to the North Atlantic is POC (ref. 14), and the present data suggest that this material is very old (on the order of 10^3 years or greater). The remaining 60%, transported as DOC (ref. 14), may be old or recently produced, depending on the region being drained. In addition, even initially young DOC may attain a considerably older mean age following modest preferential degradation of younger components by river and coastal bacteria (Fig. 1). Finally, the exchange of old organic matter adsorbed on mineral particles during mixing of river and marine waters^{24,25} will probably be an additional source of old DOC within these waters—which are then exported to the coastal and open ocean. Thus we conclude that rivers are important sources of old POC, and net sources of old

DOC, and therefore supply organic matter to the ocean margins in a pre-aged form.

The rivers sampled as part of this study are a subset of the two dominant regions of freshwater discharge to the North Atlantic: the Amazon basin (~42%) and North America (~23%)²⁹. The mean $\Delta^{14}\text{C}$ values for the Amazon and North American rivers are respectively 28‰ and 27‰ for DOC, and -145‰ and -214‰ for POC. When these are normalized for freshwater discharge, the mean values are 28‰ for DOC and -175‰ for POC. Assuming that the $\Delta^{14}\text{C}$ for DOC decreases by approximately 50‰ as a result of preferential removal of young components by bacteria during transport (Fig. 1), and using the proportions of organic carbon exported by rivers to the North Atlantic as 60% for DOC and 40% for POC (ref. 14), we calculate a net weighted mean $\Delta^{14}\text{C}$ of organic carbon exported to the North Atlantic of -84‰, or about 670 years in age.

If we assume a greater degree of bacterial utilization of young DOC components (that is, a decrease of about 200‰ in $\Delta^{14}\text{C}$ -DOC; Fig. 1b) over residence times approximating those of shelf waters (~1 yr), then the weighted mean $\Delta^{14}\text{C}$ of organic carbon (DOC plus POC) exported from rivers to the North Atlantic becomes -173‰ or 1,500 years in age. Bacteria may also preferentially utilize young components of the POC pool, making our estimates conservative.

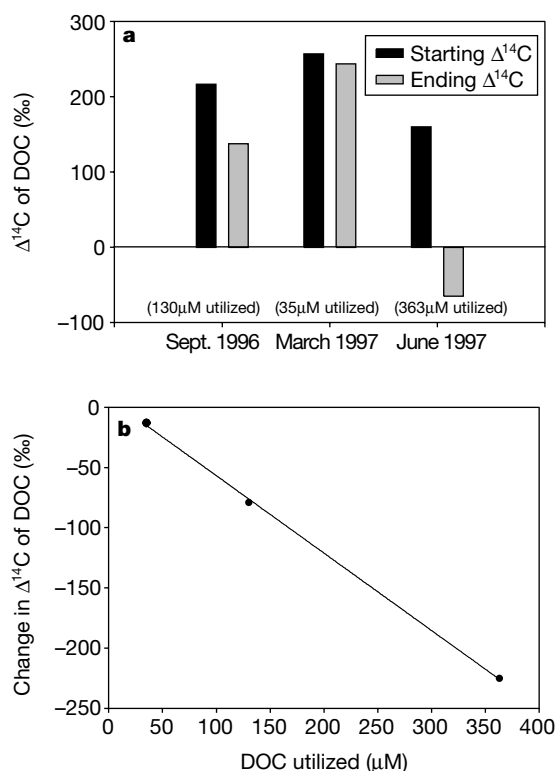


Figure 1 Results from three $\Delta^{14}\text{C}$ -DOC utilization experiments performed on water from the York River. Fresh water was filtered (0.7 μm), then incubated in the dark at *in situ* temperatures (23 °C, 10 °C and 20 °C for the September, March and June experiments, respectively). **a**, Changes in the $\Delta^{14}\text{C}$ and concentrations of DOC during incubation. The concentrations of DOC and $\Delta^{14}\text{C}$ -DOC values were measured at the beginning and end of each incubation by the methods outlined in Table 1. Values in parentheses indicate the amounts of DOC utilized by bacteria during each long-term incubation. September 1996 and March 1997 incubations both lasted for 60 d, while the June 1997 incubation lasted 365 d. The smaller DOC and $\Delta^{14}\text{C}$ changes in the March 1997 incubation were probably due to low (10 °C) incubation temperatures. **b**, The change in $\Delta^{14}\text{C}$ versus the amount of DOC utilized by bacteria for incubations in **a**. The strong relationship ($r^2 = 0.99$, $P < 0.05$) indicates that as bacteria utilize greater amounts of DOC, the $\Delta^{14}\text{C}$ of the remaining fraction becomes progressively lower.

Industrial and agricultural activities in North America and Europe would be expected to result in petroleum-derived (that is, devoid of ^{14}C) older soil organic matter being introduced to watersheds, further biasing these estimates of DOC and POC to old mean ages. In order to extend such estimates of the ages of riverine and terrestrial organic matter to a global scale, it will be necessary to sample additional rivers world-wide. This area of research would further benefit from a more rigorous attempt to characterize the mechanisms behind ageing within the POC component. Even so, the present data show that the ages of DOC and POC exported to the North Atlantic are on average much older and more variable than previously assumed; the same may well be true for river water exported to other ocean waters. □

Received 6 July; accepted 8 November 2000.

- Hedges, J. I., Keil, R. G. & Benner, R. What happens to terrestrial organic matter in the ocean? *Org. Geochem.* **27**, 195–212 (1997).
- Meybeck, M. Riverine transport of atmospheric carbon: sources, global typology, and budget. *Wat. Air Soil Pollut.* **70**, 443–463 (1993).
- Bauer, J. E., Williams, P. M. & Druffel, E. R. M. ^{14}C activity of dissolved organic carbon fractions in the North Central Pacific and Sargasso Sea. *Nature* **357**, 667–670 (1992).
- Druffel, E. R. M., Williams, P. M., Bauer, J. E. & Ertel, J. Cycling of dissolved and particulate organic matter in the open ocean. *J. Geophys. Res.* **97**, 15639–15659 (1992).
- Williams, P. M. & Druffel, E. R. M. Radiocarbon in dissolved organic carbon in the central north Pacific Ocean. *Nature* **330**, 246–248 (1987).
- Meyers-Schulte, K. J. & Hedges, J. I. Molecular evidence for a terrestrial component of organic matter dissolved in ocean water. *Nature* **321**, 61–63 (1986).
- Hedges, J. I., Hatcher, P. G., Ertel, J. R. & Meyers-Schulte, K. J. A comparison of dissolved humic substances from seawater with Amazon River counterparts by ^{13}C -NMR spectrometry. *Geochim. Cosmochim. Acta* **56**, 1753–1757 (1992).
- Opsahl, S. & Benner, R. Distribution and cycling of terrigenous dissolved organic matter in the ocean. *Nature* **386**, 480–482 (1997).
- Hedges, J. I. *et al.* Organic carbon-14 in the Amazon River system. *Science* **231**, 1129–1131 (1986).
- Spiker, E. in *Flux of Organic Carbon by Rivers to the Oceans* 79–109 (Report CONF-8009140, US Dept of Energy, Springfield, 1981).
- Kao, S. & Liu, K. Particulate organic carbon export from a subtropical mountainous river (Lanyang Hsi) in Taiwan. *Limnol. Oceanogr.* **41**, 1749–1757 (1996).
- Levin, I. & Kromer, B. Twenty years of atmospheric $^{14}\text{CO}_2$ observations at Schauinsland station, Germany. *Radiocarbon* **39**, 205–218 (1997).
- Hedges, J. I. *et al.* Composition and fluxes of particulate organic material in the Amazon River. *Limnol. Oceanogr.* **31**, 717–738 (1986).
- Ludwig, W., Probst, J.-L. & Kempe, S. Predicting the oceanic input of organic carbon by continental erosion. *Glob. Biogeochem. Cycles* **10**, 23–41 (1996).
- Richter, D. D., Markewitz, D., Trumbore, S. E. & Wells, C. G. Rapid accumulation and turnover of soil carbon in a re-establishing forest. *Nature* **400**, 56–58 (1999).
- O'Brien, B. J., Stout, J. D. & Goh, K. M. in *Flux of Organic Carbon by Rivers to the Oceans* 46–75 (Report CONF-8009140, US Dept of Energy, Springfield, 1981).
- Trumbore, S. E., Vogel, J. S. & Southon, J. R. AMS ^{14}C measurements of fractionated soil organic matter: An approach to deciphering the soil carbon cycle. *Radiocarbon* **31**, 644–654 (1989).
- Trumbore, S. E., Schiff, S. L., Aravena, R. & Elgood, R. Sources and transformation of dissolved organic carbon in the Harp Lake forested catchment: The role of soils. *Radiocarbon* **34**, 636–635 (1992).
- Degens, E. T. in *Perspectives on Biogeochemistry* (ed. Degens, E. T.) Ch. 11, 303–304 (Springer, New York, 1989).
- Coch, N. K. & Bokuniewicz, H. J. Oceanographic and geologic framework of the Hudson system. *Northeast. Geol.* **8**, 96–108 (1986).
- Stallard, R. F. & Edmond, J. M. Geochemistry of the Amazon 2. The influence of geology and weathering environment on the dissolved load. *J. Geophys. Res.* **88**, 9671–9688 (1983).
- Meade, R. H., Dunne, T., Richey, J. E., Santos, U. M. & Salati, E. Storage and remobilization of suspended sediment in the lower Amazon River of Brazil. *Science* **228**, 488–490 (1985).
- Newbold, J. D., Mulholland, P. J., Elwood, J. W. & O'Neill, R. V. Organic carbon spiralling in stream ecosystems. *Oikos* **38**, 266–272 (1982).
- Hedges, J. I. & Keil, R. G. Sedimentary organic matter preservation: an assessment and speculative synthesis. *Mar. Chem.* **49**, 81–115 (1995).
- Keil, R. G., Mayer, L. M., Quay, P. D., Richey, J. E. & Hedges, J. I. Loss of organic matter from riverine particles in deltas. *Geochim. Cosmochim. Acta* **61**, 1507–1511 (1997).
- Howarth, R. W., Fruci, J. R. & Sherman, D. Inputs of sediment and carbon to an estuarine ecosystem: influence of land use. *Ecol. Appl.* **1**, 27–39 (1991).
- Raymond, P. A. & Bauer, J. E. Bacterial consumption of DOC during transport through a temperate estuary. *Aquat. Microb. Ecol.* **22**, 1–12 (2000).
- Moran, M. A., Sheldon, W. M. & Sheldon, J. E. Biodegradation of riverine dissolved organic carbon in five estuaries of the Southeastern United States. *Estuaries* **22**, 55–64 (1999).
- Howarth, R. W. *et al.* Regional nitrogen budgets and riverine N & P fluxes for the drainages to the North Atlantic Ocean: Natural and human influences. *Biogeochemistry* **35**, 75–139 (1996).
- Vogel, J. S., Southon, J. R. & Nelson, D. E. Catalyst and binder effects in the use of filamentous graphite for AMS. *Nucl. Instrum. Methods Phys. Res. B* **29**, 50–56 (1987).
- Stuiver, M. & Polach, H. A. Discussion: reporting of ^{14}C data. *Radiocarbon* **19**, 355–363 (1977).
- Sofer, Z. Preparation of carbon dioxide for stable carbon isotope analysis of petroleum fractions. *Anal. Chem.* **52**, 1389–1391 (1980).

Acknowledgements

We thank J. Cole, N. Caraco and C. Hopkinson for help in collecting samples from the Hudson and Parker rivers; D. Wolgast for assistance in oxidizing DOC samples; M. Kashgarian, J. Southon and B. Frantz for ^{14}C analyses at the Center for AMS at Lawrence Livermore National Laboratory (LLNL); E. Druffel and S. Griffin for analysing the Amazon POC sample; J. Hobbie for comments on the manuscript; and E. Franks for $\delta^{13}\text{C}$ analyses at Woods Hole Oceanographic Institution. This work was supported by the Ocean Margins Program of the US Department of Energy, the Chemical Oceanography and Long-Term Ecological Research Programs of the US NSF, and the Center for AMS at LLNL.

Correspondence and requests for materials should be addressed to P.A.R. (e-mail: praymond@mbi.edu).

Geochemical evidence for the melting of subducting oceanic lithosphere at plate edges

G. M. Yogodzinski*, J. M. Lees†, T. G. Churikova‡, F. Dorendorf§, G. Wöerner§ & O. N. Volynets||

* Department of Geology, Dickinson College, Carlisle, Pennsylvania 17013-2896, USA

† Department of Geological Sciences, University of North Carolina, Chapel Hill, North Carolina 27516, USA

‡ Institute of Volcanic Geology and Geochemistry, Petropavlovsk-Kamchatsky, 683006 Russia

§ Geochemisches Institut, Goldschmidstrasse, 1, 37077 Göttingen, Germany

|| Deceased

Most island-arc magmatism appears to result from the lowering of the melting point of peridotite within the wedge of mantle above subducting slabs owing to the introduction of fluids from the dehydration of subducting oceanic crust¹. Volcanic rocks interpreted to contain a component of melt (not just a fluid) from the subducting slab itself are uncommon, but possible examples have been recognized in the Aleutian islands, Baja California, Patagonia and elsewhere^{2–4}. The geochemically distinctive rocks from these areas, termed 'adakites', are often associated with subducting plates that are young and warm, and therefore thought to be more prone to melting⁵. But the subducting lithosphere in some adakite locations (such as the Aleutian islands) appears to be too old and hence too cold to melt^{6,7}. This implies either that our interpretation of adakite geochemistry is incorrect, or that our understanding of the tectonic context of adakites is incomplete. Here we present geochemical data from the Kamchatka peninsula and the Aleutian islands that reaffirms the slab-melt interpretation of adakites², but in the tectonic context of the exposure to mantle flow around the edge of a torn subducting plate. We conclude that adakites are likely to form whenever the edge of a subducting plate is warmed or ablated by mantle flow. The use of adakites as tracers for such plate geometry may improve our understanding of magma genesis and thermal structure in a variety of subduction-zone environments.

In the northwesternmost Pacific, there is a dramatic shift in subduction dynamics where the Aleutian and Kamchatka arcs meet at an angle of nearly 90° (Fig. 1). Well defined, deep seismic zones indicate that the subducting Pacific plate dips to the west beneath Kamchatka, and to the north beneath the central Aleutians, but the nature of subduction beneath the junction of these arcs is not well understood. One recent study suggests that the subducting Pacific

plate is not present beneath the western Aleutians, and therefore does not bend sharply beneath the corner formed by the Aleutian–Kamchatka junction⁸. This interpretation, which is a significant departure from some previous views⁹, implies that the western Aleutians are a simple transform boundary, and that there is a window through the subducting Pacific plate beneath the western Aleutians and the Aleutian–Kamchatka junction. In turn, this implies that as the Pacific plate passes beneath the Aleutian–Kamchatka junction, it is exposed to shearing and mantle flow around its northern edge—a situation that is likely to cause this edge to be anomalously hot and/or physically ablated⁸. Northward shoaling of seismicity beneath Kamchatka supports this view of slab geometry (Fig. 2). The implications for magma genesis in the Aleutian–Kamchatka junction are clear: if the edge of the slab is heated sufficiently, it may contribute a melt (not just a fluid) to the source of magmas erupted in the vicinity of Sheveluch volcano, the northernmost active strato-volcano in Kamchatka (Fig. 1).

The effects of slab melting on arc magma geochemistry are incompletely understood, but many workers believe that a distinctive class of magnesian and calc-alkaline andesites and dacites termed 'adakites'⁵ (see Fig. 3 regarding our use of this term) may be examples of arc volcanic rocks that are produced in part by melting of subducting oceanic crust. Studies of adakites indicate that the geochemical consequences of slab melting may be anomalously steep rare-earth-element patterns, high Sr contents and high Sr/Y ratios in andesitic-to-dacitic rocks that may also be relatively rich in compatible elements such as Mg, Cr and Ni. These features, which are most commonly seen in arcs where the subducting oceanic lithosphere is relatively young⁵, are thought to be produced when subducting oceanic crust melts at high pressures to produce silicic magmas that rise through—and react with—the overlying wedge of mantle peridotite².

Figure 3 shows that volcanic rocks of the Sheveluch area are significantly more Mg-rich (lower FeO^*/MgO) at a given SiO_2 content than are rocks from the adjacent volcanoes of the Klyuchevskoy group. (FeO^* is total iron content calculated as FeO .) Figure 4 shows that with regard to Sr/Y, the Sheveluch rocks fall

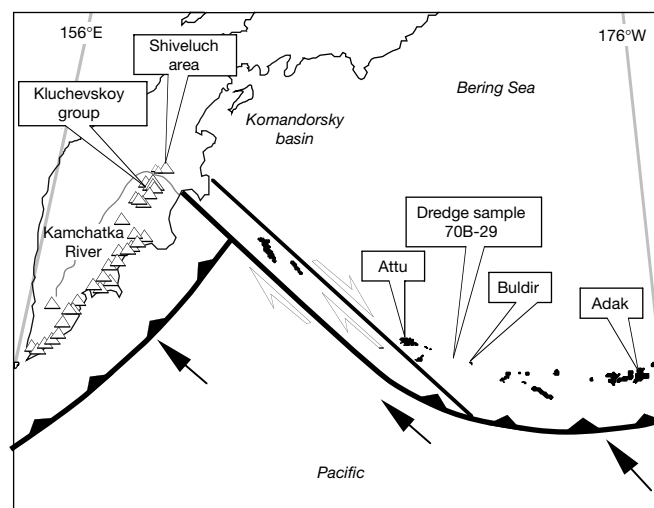


Figure 1 Map view of the study area. The figure shows the oblique subduction zone of the central Aleutians (near Adak), the transform-type boundary in the westernmost Aleutians, and the sharp bend into the subduction system of the Kurile–Kamchatka arc. The Klyuchevskoy volcanic group and the Sheveluch area mark the location of the Kamchatka central depression. Exceptionally voluminous arc volcanism in the central depression, and the slight offset of the magmatic front in this area (apparently due to a slight flattening of the slab dip²⁷) are believed to be evidence of mantle flow around the northern edge of the subducting Pacific plate as it passes beneath the Aleutian–Kamchatka junction⁸.

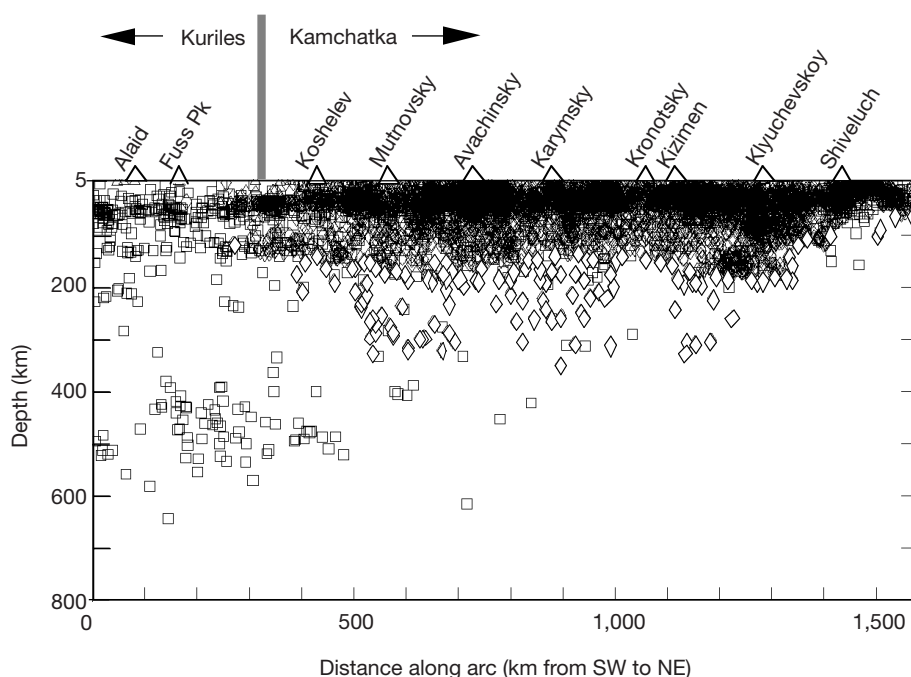


Figure 2 Along-arc seismicity in the Kurile–Kamchatka system. The figure illustrates the northward shoaling of seismicity towards the Aleutian–Kamchatka junction. The absence of deep earthquakes beneath the Aleutian–Kamchatka junction is interpreted to imply that mantle flow around the northern edge of the Pacific plate (possibly induced by slab roll-back) has significantly warmed and physically ablated the subducting plate beneath this

junction⁸. Distinctive geochemical features of volcanic rocks from the Sheveluch area (Figs 3, 4) are interpreted to be the result of anomalous heating of the plate along its northern edge beneath the Aleutian–Kamchatka junction. Earthquake locations are from refs 27 (diamonds) and 28 (squares). White triangles mark the locations of the major volcanic centres along the arc.

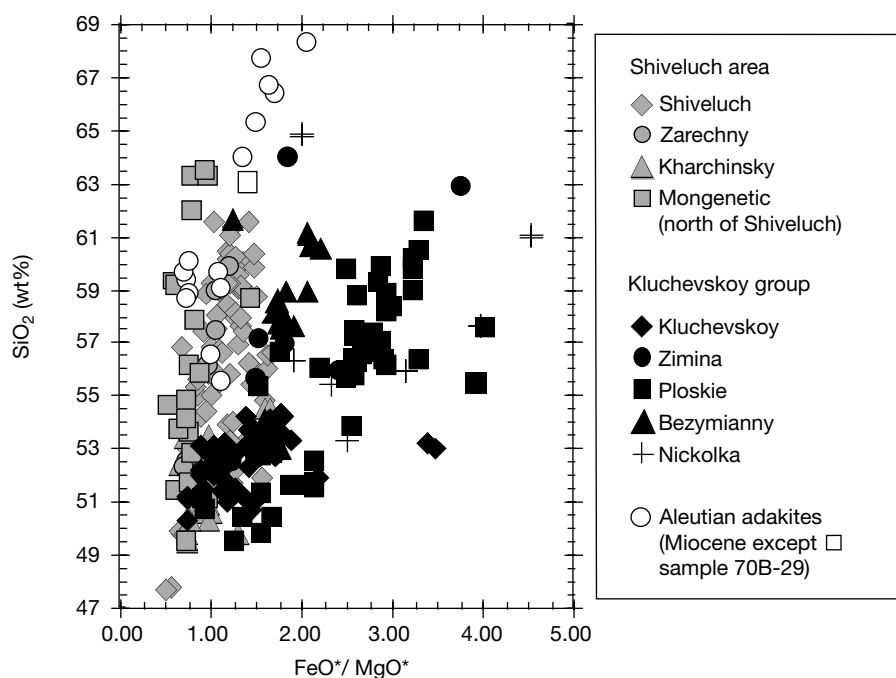


Figure 3 Whole-rock SiO_2 content versus FeO^*/MgO ratio in volcanic rocks of the Kamchatka central depression compared to Aleutian adakites. Kamchatka data are grouped geographically from areas immediately north (Sheveluch area) and south (Klyuchevskoy group) of the Kamchatka River (see also Fig. 1). We use the term ‘adakite’ throughout this Letter to describe calc-alkaline andesites and dacites that are relatively Mg-rich (low FeO^*/MgO , high $\text{Mg}/(\text{Mg}+\text{Fe})$) and also have anomalously high Sr/Y (>50) compared to ‘normal’ arc volcanic rocks (Sr/Y not shown here, see Fig. 4). This graph shows that rocks from the Sheveluch area are more Mg-rich relative to SiO_2 than are those

from the adjacent Klyuchevskoy group. We attribute the contrasting trends, in part, to the presence of primitive andesites containing an adakitic geochemical component in the rocks of the Sheveluch area. The primitive andesites have high SiO_2 relative to FeO^*/MgO , and this difference is interpreted to persist throughout the fractionation and mixing processes by which these magmas evolve²⁴. This graph includes 142 data points from the Sheveluch area and 147 from the Klyuchevskoy group. Published data ($n = 74$) are from refs 29–32.

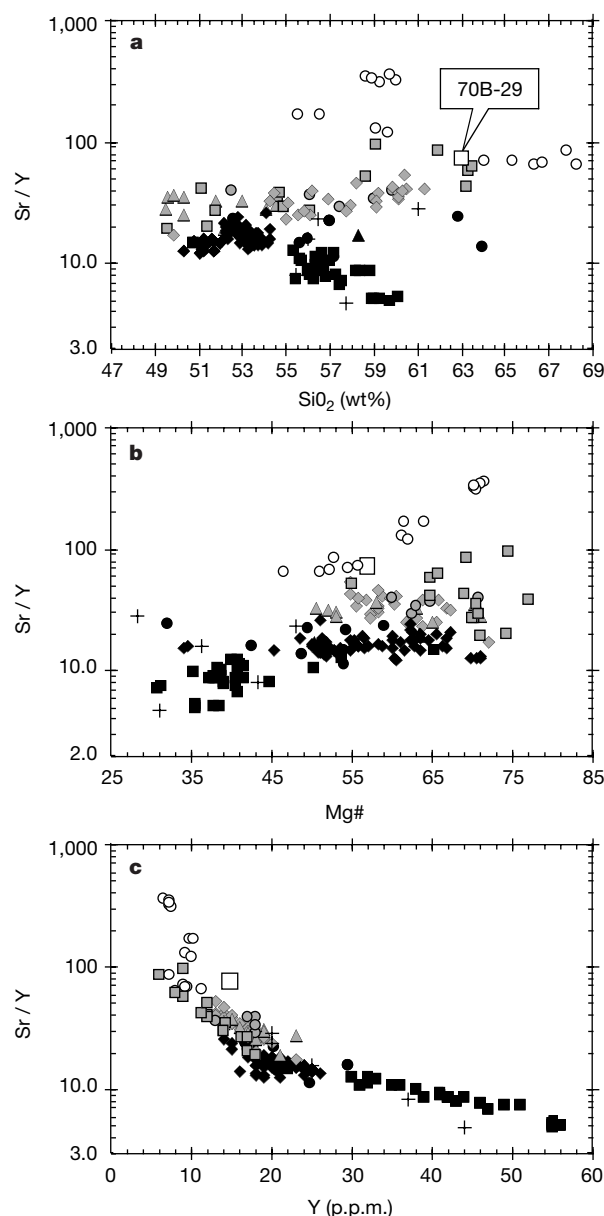


Figure 4 Sr/Y ratio versus SiO₂ content, Mg# and Y content for volcanic rocks of the Kamchatka central depression compared to Aleutian adakites. Symbols are the same as in Fig. 3; Mg# = $100[Mg/(Mg + Fe)]$ calculated on a molar basis using total Fe content. We emphasize Sr/Y because this parameter is interpreted to be a strong indicator of slab melting, and because these trace elements are relatively well represented in the available data. The Sheveluch and Klyuchevskoy data fields overlap in the range of Sr/Y = 10–25, which are common values for primitive arc volcanic rocks. The distinctive aspect of the data set from the Sheveluch group is the common occurrence of low Y (<17 p.p.m.) with high Sr/Y (>30), especially in samples with relatively high Mg# and high SiO₂. A crystal fractionation or AFC (assimilation-fractional crystallization) origin for this trend is unlikely, though the effects of fractionation are clear in the Klyuchevskoy data which show strong trends toward decreasing Sr/Y with increasing SiO₂ (a), decreasing Mg# (b), and increasing Y (c). For some samples plotted here, including most of the adakites and all of the samples from Ziminy volcano, available Yb data have been used to estimate the concentration of Y by assuming that Y/Yb = 10.5. Variation of the assumed Y/Yb value within reasonable limits (9–12) has no significant effect on the appearance of these graphs. The distinctive characters of the samples from Sheveluch area and the Klyuchevskoy group are clearly resolved, even though data were acquired by several laboratories over many years. The number of data points plotted here differs from Fig. 3 because Sr and Y are available for only a subset of samples. Published data sources here are the same as in Fig. 3.

between Aleutian adakites and ‘normal’ arc volcanic rocks of the Klyuchevskoy group. These data show that the rocks of the Sheveluch area are chemically distinctive in ways that are consistent with the possible presence of a slab-melt component in their source. The strong adakite geochemical signature in samples from the monogenetic centres probably reflects the fact that these rocks did not pass through the magmatic plumbing system beneath Sheveluch proper, and therefore underwent comparatively little fractionation and mixing with ‘normal’ arc magmas—processes that will generally dilute the adakite geochemical component from the slab¹⁰.

Combining geophysical, geochemical and dynamical observations, we find significant support for the idea of a slab window and subducting plate edge beneath the Aleutian–Kamchatka junction⁸. This view implies that there exists an active tear in the subducting Pacific plate somewhere beneath the western Aleutians (Fig. 5). The location of the tear is not constrained by seismicity, but the geochemistry of active western Aleutian magmatism suggests that it may be located between Buldir and Attu islands, where the arc switches from convergent to strike-slip motion. Sampling of volcanic rocks in this area is sparse, but in this context a singular rock, sample 70B-29, is of interest. Sample 70B-29 is a hornblende-bearing andesite with geochemistry that places it outside the norm for Aleutian volcanic rocks, but squarely among the adakites (Figs 3, 4). Sample 70B-29 is by far the youngest Aleutian adakite available (<610,000 years old¹¹), and it is the only one that can be placed with confidence in the modern tectonic context. This sample was dredged from the sea floor approximately 80 km west of Buldir island¹¹, where the subducting plate is approximately 50 Myr old¹². According to thermal models, this subducting lithosphere is too old and cold to melt beneath the arc^{6,7}, so the association of adakitic volcanism with young subducting lithosphere cannot apply in this case. We argue that adakite sample 70B-29 was produced by melting along the edge of the Pacific plate at the point where it is being torn in an unzipping motion as it passes beneath the western Aleutians (Fig. 5). We infer from this case that Miocene adakites in the Aleutians probably also resulted from melting along the edge of a torn Pacific plate, possibly as the locus of the tear migrated along the arc during Neogene times.

In general, it seems that melting at plate edges is possible in all cases of ridge subduction and slab window formation (for example, Baja, Patagonia). A slab window similar that of the Aleutian–Kamchatka junction exists in northern California, where the Mendocino triple junction bounds the Cascade subduction system to the north, and the San Andreas transform to the south. In the Cascades, adakitic andesites occur only at Mts Shasta and Lassen^{13,14}, which lie close to the triple junction, above the southern edges of the subducting Juan de Fuca and Gorda plates. Here, the adakite geochemical component seems to be spatially linked to plate edges near the triple junction, but apparently not to the subduction of young lithosphere beneath Oregon and southern Washington^{15,16}. The occurrence of geochemically distinctive primitive andesites in the Pliocene-age Clear Lake Volcanics (high Mg content and La/Yb ratio, low ⁸⁷Sr/⁸⁶Sr ratio)^{17,18} suggests that adakite geochemistry may also be a common feature of magmatism associated with the northward migration of the Mendocino triple junction in the Neogene. In Costa Rica, adakite genesis has been linked to the melting of young and hot lithosphere during subduction of an extinct segment of the Galapagos spreading centre¹⁹, but if adakites can also form by melting along the edges of plates that are older, then the possible role of alternative subduction geometries must be considered²⁰. The unusual case of adakite genesis at Solander island, located south of New Zealand’s South Island²¹ may also be associated with a subducting plate edge. The tectonic geometry in this place is poorly understood²¹, but geophysical observations are consistent with the bending, and possible tearing, of the Australian plate as it passes beneath the Pacific plate at a highly oblique angle²².

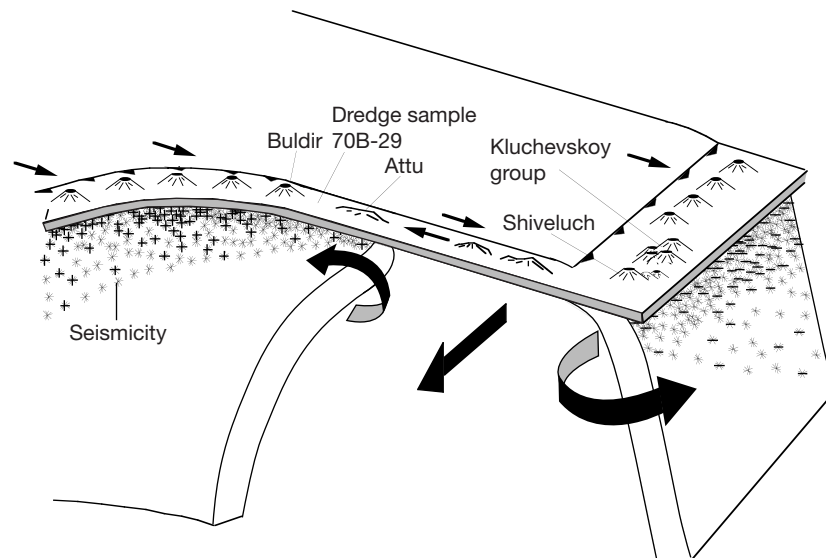


Figure 5 Perspective drawing showing a torn Pacific plate subducting to the north beneath the central Aleutians (Adak) and to the west beneath Kamchatka. The geometry shown here implies that the Pacific plate is being torn in an unzipping motion as the Aleutian slab sinks to the north beneath the Bering Sea. Adakitic volcanism is observed at Sheveluch volcano, immediately above the torn plate edge beneath Kamchatka, and above the inferred position of the active tear in the Pacific plate beneath the western Aleutians (location of andesite sample 70-B29). The large arrows indicate

asthenospheric flow around the plate edges and through the slab window. This kind of mantle flow would explain the melting of relatively old subducting plates along their edges, as well as the relatively high heat flow from the area of the Komandorsky basin in the western Bering Sea (location in Fig. 1). Asterisks indicate the seismicity (schematically) which shoals to the north beneath Kamchatka (Fig. 2) and to the west beneath the Aleutians.

In this case we expect adakite genesis through melting of the leading edge of the Australian plate as it enters warm asthenosphere beneath the Pacific plate.

The Aleutian, Kamchatkan and other examples cited here emphasize that adakite genesis is linked to unusual conditions in the subducting plate that are produced not only by slab age⁵, but also by slab geometry. The wide geochemical variation in the western Aleutian–Kamchatka region (Figs 3, 4) suggests that in general, arc magmatism may exist on a continuum between relatively hot mantle systems, which are dominated by primary basalt (for example, Klyuchevskoy), and hot slab systems, which are dominated by relatively cool and hydrous primary andesite (for example, Sheveluch). In this view, adakites with Mg number (Mg#) greater than 60 (Fig. 4b) represent the extreme case of primitive arc magmatism in a hot slab system. Relatively low magmatic output of such systems (for example, western Aleutians^{23,24}, southwest Japan^{25,26}) may result from dehydration of the slab at shallow depths²⁶, or from the cooling of the mantle wedge by slab-derived melts. Endmember hot-slab and hot-mantle systems will generally also show contrasting seismicity and seismic structure²⁶. This view of arc magmatism underlines the need to improve our understanding of subduction dynamics and the physical conditions of the slab and mantle wedge, and their connection to primary magma geochemistry in a variety of subduction settings. □

Received 27 June; accepted 13 November 2000.

- Perfit, M. R., Gust, D. A., Bence, A. E., Arculus, R. J. & Taylor, S. R. Chemical characteristics of island arc basalts: implications for mantle sources. *Chem. Geol.* **30**, 227–256 (1980).
- Kay, R. W. Aleutian magmatic andesites: melts from subducted Pacific ocean crust. *J. Volcanol. Geotherm. Res.* **4**, 117–132 (1978).
- Rogers, G., Saunders, A. D., Terrell, D. J., Verma, S. P. & Marriner, G. F. Geochemistry of Holocene volcanic rocks associated with ridge subduction in Baja, California, Mexico. *Nature* **315**, 389–392 (1985).
- Stern, C. R., Futa, K. & Muehlenbachs, K. in *Andean Magmatism: Chemical and Isotopic Constraints* (eds Harmon, R. S. & Barreiro, B. A.) 31–47 (Shiva, Nantwich, UK, 1984).
- Defant, M. J. & Drummond, M. S. Derivation of some modern arc magmas by melting of young subducted lithosphere. *Nature* **347**, 662–665 (1990).
- Davies, J. H. & Stevenson, D. J. Physical model of source region of subduction zone volcanics. *J. Geophys. Res.* **97**, 2037–2070 (1992).

- Peacock, S. M., Rushmer, T. & Thompson, A. B. Partial melting of subducting oceanic crust. *Earth Planet. Sci. Lett.* **121**, 227–244 (1994).
- Davaille, A. & Lees, J. M. Thermal modeling of subducted plates: tear and hot spot at the Kamchatka corner. *Geophys. Res. Lett.* (in the press).
- Creager, K. C. & Boyd, T. M. The geometry of Aleutian subduction: three dimensional kinematic flow model. *J. Geophys. Res.* **96**, 2293–2307 (1991).
- Yogodzinski, G. M. & Kelemen, P. B. Slab melting in the Aleutians: implications of an ion probe study of clinopyroxene in primitive adakite and basalt. *Earth Planet. Sci. Lett.* **158**, 53–65 (1998).
- Scholl, D. W., Marlow, M. S., MacLeod, N. S. & Buffington, E. C. Episodic Aleutian Ridge igneous activity: implications of Miocene and younger submarine volcanism west of Buldir Island. *Geol. Soc. Am. Bull.* **87**, 547–554 (1976).
- Lonsdale, P. Paleogene history of the Kula plate: offshore evidence and onshore implications. *Geol. Soc. Am. Bull.* **100**, 733–754 (1988).
- Baker, M. B., Grove, T. L. & Price, R. Primitive basalts and andesites from the Mt. Shasta region, N. California: products of varying melt fraction and water content. *Contrib. Mineral. Petrol.* **118**, 111–129 (1994).
- Borg, L. E., Clynnne, M. A. & Bullen, T. D. The variable role of slab-derived fluids in the generation of a suite of primitive calc-alkaline lavas from the southernmost Cascades, California. *Can. Mineral.* **35**, 425–452 (1997).
- Conrey, R. M. Diverse primitive magmas in the Cascade arc, northern Oregon and southern Washington. *Can. Mineral.* **35**, 367–396 (1997).
- Bacon, C. R. et al. Primitive magmas at five Cascade volcanic fields: melts from hot, heterogeneous sub-arc mantle. *Can. Mineral.* **35**, 397–423 (1997).
- Johnson, C. M. & O'Neil, J. R. Triple junction magmatism: a geochemical study of Neogene volcanic rocks in western California. *Earth Planet. Sci. Lett.* **71**, 241–262 (1984).
- Hearn, B. C., Donnelly-Nolan, J. M. & Goff, F. E. in *Research in the Geysers-Clear Lake Geothermal Area, Northern California* (eds McLaughlin, R. J. & Donnelly-Nolan, J. M.) 25–45 (US Government Printing Office, Washington DC, 1981).
- Johnston, S. T. & Thorkelson, D. J. Cocos-Nazca slab window beneath Central America. *Earth Planet. Sci. Lett.* **146**, 465–474 (1997).
- Abratis, M. & Worner, G. Ridge collision, slab window formation and the flux of Pacific asthenosphere into the Caribbean realm in southern Costa Rica. *Geology* (in the press).
- Reay, A. & Parkinson, D. Adakites from Solander Island, New Zealand. *NZ J. Geol. Geophys.* **40**, 121–126 (1997).
- Furlong, K. P. & Anderson, H. Lithospheric tectonics of a transpressional plate boundary: Fiordland, New Zealand [abstract]. *Ann. Geophys.* **16**, 45 (1988).
- Yogodzinski, G. M., Volynets, O. N., Koloskov, A. V., Seliverstov, N. I. & Matvenkov, V. V. Magnesian andesites and the subduction component in a strongly calc-alkaline series at Piip Volcano, Far Western Aleutians. *J. Petrol.* **35**, 163–204 (1994).
- Yogodzinski, G. M., Kay, R. W., Volynets, O. N., Koloskov, A. V. & Kay, S. M. Magnesian andesite in the western Aleutian Komandorsky region: Implications for slab melting and processes in the mantle wedge. *Geol. Soc. Am. Bull.* **107**, 505–519 (1995).
- Morris, P. A. Slab melting as an explanation of Quaternary volcanism and aseismicity in southwest Japan. *Geology* **23**, 395–398 (1995).
- Peacock, S. M. & Wang, K. Seismic consequences of warm versus cool subduction metamorphism:

- examples from southwest and northeast Japan. *Science* **286**, 937–939 (1999).
27. Gorbato, A., Kostoglodov, V., Gerardo, S. & Gordeev, E. Seismicity and structure of the Kamchatka subduction zone. *J. Geophys. Res.* **102**, 17883–17898 (1997).
 28. Engdahl, E. R., van der Hilst, R., Kirby, S. H. & Ekstrom, G. A global survey of slab structures and internal processes using a combined data base of high-resolution earthquake hypocenters, tomographic images and focal mechanism data. *Seismol. Res. Lett.* **69**, 153–154 (1998).
 29. Hochstaedter, A. G., Kepezhinskas, P., Defant, M., Drummond, M. & Koloskov, A. Insights into the volcanic arc mantle wedge from magnesian lavas from the Kamchatka arc. *J. Geophys. Res.* **101**, 697–712 (1996).
 30. Kepezhinskas, P. et al. Trace element and Sr–Nd–Pb isotopic constraints on a three-component model of Kamchatka arc petrogenesis. *Geochim. Cosmochim. Acta* **61**, 577–600 (1997).
 31. Kersting, A. B. & Arculus, R. J. Klyuchevskoy Volcano, Kamchatka, Russia: the role of high-flux, recharged, tapped and fractionated magma chamber(s) in the genesis of high- Al_2O_3 from high-MgO basalt. *J. Petrol.* **35**, 1–42 (1994).
 32. Ozerov, A. Y., Ariskin, A. A., Kyle, P., Bogoyavlenskaya, G. E. & Karpenko, S. F. Petrological-geochemical model for genetic relationships between basaltic and andesitic magmatism of Klyuchenskoi and Bezmyanniy volcanoes, Kamchatka. *Petrology* **5**, 550–569 (1997).

Acknowledgements

We thank P. Kelemen and K. Furlong for discussions; J. Morris and M. Defant for comments on the manuscript; and A. Bellousov, M. Bellousova, M. Ejzak, A. Koloskov, G. Ponomarev and V. Ponomareva for assistance in the field. This work was supported by the US NSF (G.M.Y. and J.M.L.), the German Science Foundation, the Volkswagen Foundation and the European Union (INTAS) (to G.W. and T.C.), the Russian Foundation for Basic Research (O.V. and T.C.), and by a grant from the Whitaker Foundation to Dickinson College.

Correspondence and requests for materials should be addressed to G.M.Y. (e-mail: yogodzin@dickinson.edu).

A bizarre predatory dinosaur from the Late Cretaceous of Madagascar

Scott D. Sampson*, Matthew T. Carrano† & Catherine A. Forster†

* Utah Museum of Natural History and Department of Geology and Geophysics, University of Utah, 1390 East Presidents Circle, Salt Lake City, Utah 84112-0050, USA

† Department of Anatomical Sciences, Health Sciences Center, State University of New York, Stony Brook, New York 11794-8081, USA

Here we report the discovery of a small-bodied (~1.8 m) predatory dinosaur from the Late Cretaceous (Maastrichtian) of Madagascar. *Masiakasaurus knopfleri*, gen. et sp. nov., represented by several skull elements and much of the postcranial skeleton, is unique in being the only known theropod with a highly procumbent and distinctly heterodont lower dentition. Such a derived dental morphology is otherwise unknown among dinosaurs. Numerous skeletal characteristics indicate that *Masiakasaurus* is a member of Abelisauroidae, an enigmatic clade of Gondwanan theropods. Previously, small-bodied abelisauroids were known only from Argentina^{1–3}. The occurrence of *Masiakasaurus* on Madagascar suggests that small-bodied abelisauroids, like their larger-bodied counterparts, were more cosmopolitan, radiating throughout much of Gondwana and paralleling the diversification of small coelurosaur theropods in Laurasia.

Several expeditions^{4–7} have recovered abundant, well-preserved skeletal remains of dinosaurs and other vertebrates from the Upper Cretaceous (Maastrichtian) Maevarano Formation of northwestern Madagascar. The non-avian dinosaur fauna includes a large abelisauroid theropod, *Majungatholus atopus*⁵, and at least two titanosaurian sauropods. Remains of at least five species of birds have also been recovered, including the basal avian *Rahonavis ostromi*⁷. Our study describes a previously unknown small-bodied form recovered from the same deposits.

Saurischia Seeley 1888

Theropoda Marsh 1881

Abelisauroidae Bonaparte 1991

Masiakasaurus knopfleri gen. et sp. nov.

Etymology. From *masiaka* (Malagasy, meaning vicious), *sauros* (Greek, meaning lizard) and *knopfleri* (after singer/songwriter Mark Knopfler, whose music inspired expedition crews).

Holotype. Université d'Antananarivo (UA) 8680, well-preserved right dentary with several teeth (Fig. 1).

Referred specimens. Field Museum of Natural History (FMNH PR 2108–2182): maxilla; dentaries; splenial; cervical, dorsal, sacral and caudal vertebrae; dorsal rib; humeri; manual phalanges and ungual; pubes; femora; tibiae; tibia/fibula/astragalocalcaneum; metatarsals II and III; pedal phalanges and unguals. UA 8681–8696: dentary; cervical, dorsal, and caudal vertebrae; femora; tibia/astragalocalcaneum; pedal phalanges; and unguals.

Localities and horizon. All specimens are from the Anembalemba Member of the Upper Cretaceous (Maastrichtian) Maevarano Formation, Mahajanga Basin, near the village of Berivotra, northwestern Madagascar⁸. With few exceptions, elements attributable to *Masiakasaurus*, including the holotype dentary (UA 8680), were recovered as isolated specimens from a 3-m² area in one stratigraphic horizon of a single locality, MAD 93-18.

Diagnosis. Differs from all known theropods in that the four dentary teeth most rostral in position are procumbent, with the first tooth set in a large, ventrally expanded alveolus that is almost horizontal in orientation. Also differs from all known theropods in that it has a strongly heterodont lower dentition: the first four teeth are elongate and weakly serrated, with labiolingually positioned carinae. Each of these four teeth terminates in a pointed apex that hooks caudally. The teeth become increasingly recurved and transversely compressed with increasing caudal position in the jaw, and possess more standard, mesiodistally positioned carinae.

Description. Together, the specimens referred to *Masiakasaurus knopfleri* account for about 40% of the skeleton (Fig. 1). The concentration of isolated *Masiakasaurus* elements at MAD 93-18 includes remains of at least six individuals. All of these specimens are assigned to a single species, despite many of the elements (for example, femora, tibiae and vertebrae) being represented by several specimens, as none shows evidence of belonging to more than one taxon of small-bodied non-avian theropod. Two osteological features (closure of the vertebral sutures and fusion of the crural and tarsal elements) indicate that the largest materials represent adult or near-adult individuals of this small-bodied taxon.

The maxilla, which is represented by a single partial specimen (Fig. 1; FMNH PR 2183), has a pronounced, raised external rim forming the perimeter of the extensive antorbital fossa. Although lacking accessory foramina, a deep accessory fossa occupies the rostral portion of the antorbital fossa. Seven alveoli are preserved and, although a small portion of the caudal alveolar margin is missing, it is unlikely that the maxillary tooth count exceeded ten. Unerupted teeth, which are preserved in the third and fifth alveoli, are transversely compressed, recurved and bear moderate serrations. Although the first tooth is absent—and thus the crown orientation cannot be determined with confidence—the first alveolus is oriented at about 45° to the horizontal, indicating that the rostrally positioned teeth in the upper jaw may have been procumbent.

The specialized dentary (Fig. 2), represented by four specimens, has a greatly emarginated caudal end that indicates an enlarged intramandibular fenestra, as in Abelisauridae. The number of tooth positions in the dentary of *Masiakasaurus* varies from 10 to 12. The slightly enlarged first alveolus is oriented almost horizontally to such an extent that the first tooth is directed forward. The first four teeth are radially arrayed, and become progressively more vertical and parasagittal with increased caudal position.

The bizarre morphologies of these first four teeth are unique among theropods. The caudal carina is labial (lateral), whereas the

rostral carina has migrated to a lingual (medial) position, resulting in a tooth with a convex rostral face and a slightly concave caudal face. The carinae of these teeth converge to a sharp, caudolaterally hooked tip, and a series of longitudinal grooves cover the caudal surface of the tooth between the carinae. Further back in the dentary, the teeth gradually assume a more typical theropod form, being recurved and transversely compressed, as the carinae migrate into more rostral and caudal positions. Although heterodonty has been described for other theropod taxa (for example, ref. 9), the degree of dental differentiation does not approach that of *Masiakasaurus*. Modern mammals with procumbent dentitions, such as some caenolestid marsupials, are often insectivorous,

using these teeth to assist in procuring prey items. Although the specific diet of *Masiakasaurus* remains speculative, such a deviation in dental morphology probably indicates a marked divergence from the typical theropod diet.

The cervical vertebrae are moderately elongate and lack any vertical offset of the cranial face over the caudal face. The centra are broad, rather than tall, with no obvious pneumatic features. In contrast, the neural arches are heavily pneumatized. The neural spines are cranially positioned, greatly abbreviated in length, and are short dorsoventrally—exceeded in height by the postzygapophyses. The postzygapophyses are elongate and swept back, forming a deep, V-shaped notch that terminates cranially at the neural spine.

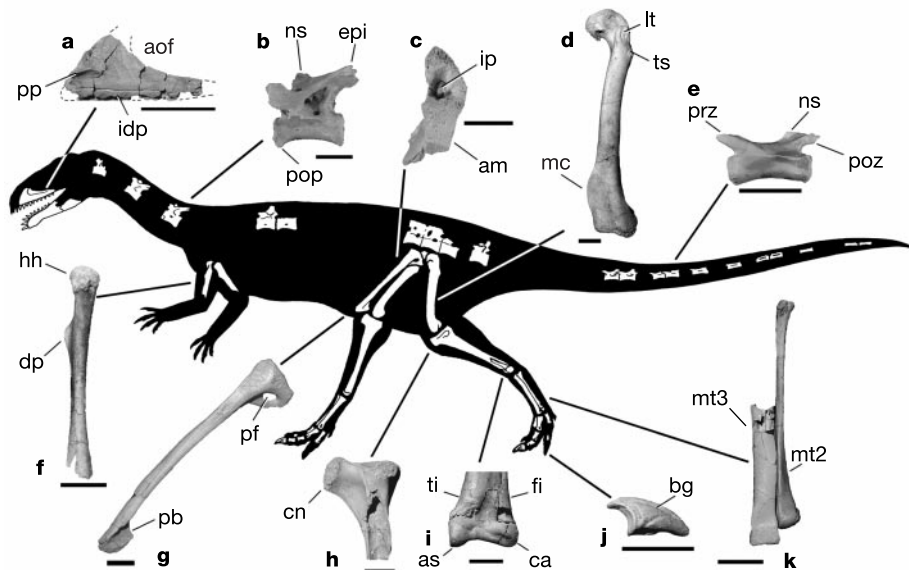


Figure 1 Skeletal anatomy of *Masiakasaurus knopfleri*, based on a composite of isolated specimens from locality MAD 93-18. **a**, Partial right maxilla (FMNH PR 2183) in medial view. Dashed lines indicate reconstructed outline of the element. **b**, Posterior cervical vertebra (FMNH PR 2139) in left lateral view. **c**, Left pubis (FMNH PR 2108) in dorsal view. **d**, Left femur (FMNH PR 2117) in cranial view. **e**, Medial caudal vertebra (FMNH PR 2126) in left lateral view. **f**, Partial right humerus (FMNH PR 2143) in medial view. **g**, Left pubis (FMNH PR 2108) in lateral view. **h**, Proximal left tibia (FMNH PR 2118) in lateral view. **i**, Left tibia/fibula/astragalocalcaneum (FMNH PR 2112) in cranial view. Dashed lines

emphasize contacts between individual elements. **j**, Pedal ungual (FMNH PR 2155) in lateral view. **k**, Right metatarsals II (FMNH PR 2129) and III (FMNH PR 2151) in cranial view. Scale bars, 1 cm. am, acetabular margin; aof, antorbital fossa; as, astragalus; bg, blood groove; ca, calcaneum; cn, cnemial crest; dp, deltopectoral crest; epi, epipophysis; fi, fibula; hh, humeral head; idp, interdental plates; ip, pit for ilio-pubic articulation; lt, lesser trochanter; mc, mediolateral crest; mt2, metatarsal II; mt3, metatarsal III; ns, neural spine; pb, pubic boot; pf, pubic fenestra; poz, postzygapophysis; pop, parapophysis; pp, palatal process; prz, prezygapophysis; ti, tibia; ts, trochanteric shelf.

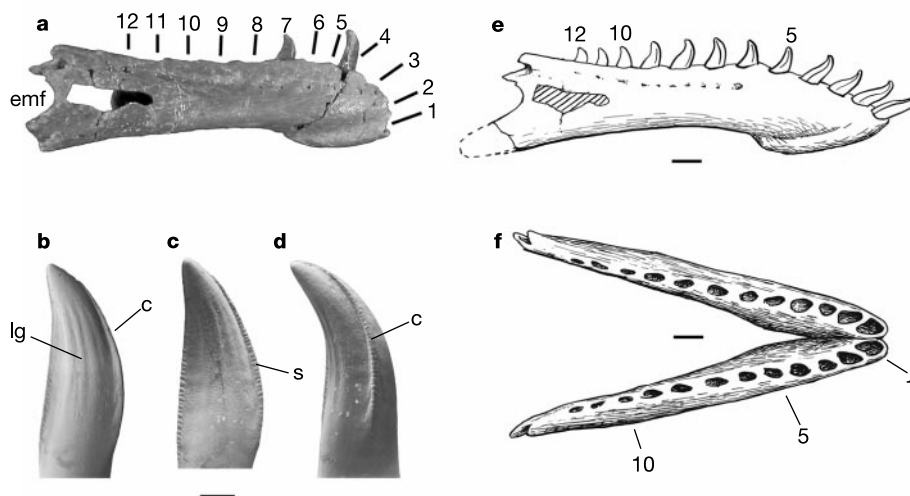


Figure 2 Dentary and lower dentition of *Masiakasaurus knopfleri*. **a**, Holotype dentary (UA 8680) in right lateral view. The tooth in position four has been displaced caudally post mortem. **b–d**, Scanning electron microscope photographs of isolated dentary teeth (FMNH uncatalogued). **b**, Anterior tooth, position 2–4 in lingual view. **c**, Anterior tooth, position 1–3, in lingual view. **d**, Anterior tooth, position 1–3, in lateral view. **e**,

Reconstructed dentary with full complement of teeth in right lateral view. **f**, Reconstructed dentary in dorsal view, showing relative sizes and orientations of alveoli. **c**, carina; emf, external mandibular fenestra; lg, longitudinal groove; s, serration. Tooth positions are numbered sequentially from cranial to caudal positions. Scale bars, 10 mm (**a**, **e–f**); 5 mm (**b–d**).

The humerus has a straight shaft, with a bulbous head that is separated from a well-developed deltopectoral crest. The femur is strongly bowed, with a greatly enlarged mediodistal crest. Proximally, the relatively long, slender tibia has a pronounced cnemial crest that curves dorsolaterally; distally, this element appears to back both the astragalus and calcaneum. The astragalus has a plate-like, rectangular ascending process that is fused with the fibula along its lateral margin. Metatarsal II is relatively straight, with the proximal two-thirds reduced to a thin shaft. The weakly curved pedal unguals possess a triangular arrangement of vascular grooves on both the lateral and medial surfaces.

Masiakasaurus can be clearly diagnosed as a theropod dinosaur on the basis of several shared, derived characters, such as recurved, serrated, laterally compressed teeth, trenchant manual unguals and markedly thin-walled skeletal elements^{10,11}. Most phylogenetic analyses of Theropoda^{10–13} have recognized two major clades, Ceratosauria (Coelophysoidea and Neoceratosauria) and Tetanurae (Spinosauroidea, Allosauroidea, Coelurosauria). However, a few studies^{14–16} have suggested that Ceratosauria is paraphyletic, with Coelophysoidea and Neoceratosauria as successive sister taxa to a monophyletic Tetanurae.

Masiakasaurus possesses a number of synapomorphies that support its membership in Abelisauroidae (Abelisauridae plus Noasauridae)¹⁷ (Fig. 3). These include: a specialized caudal dentary margin with enlarged intramandibular fenestra; straight-shafted humerus with enlarged, bulbous head; peg-and-socket articulation of ilium and pubis; mediodistal crest of femur hypertrophied to form a thin lamina; large, laterally curved and dorsally elevated cnemial crest of the tibia; and triangular arrangement of grooves on

the pedal unguals. In the Abelisauroidae, abelisaurids are all relatively large-bodied theropods that are identified by a suite of derived features, and are currently known from the Late Cretaceous of India, Madagascar and Argentina^{2,3,5,16,17}. In contrast, noasaurids are small-bodied forms previously represented by one, or possibly two, fragmentary taxa restricted to the Late Cretaceous of Argentina^{1–3}. Given that *Masiakasaurus* shares derived features with the noasaurid *Noasaurus leali* from the Late Cretaceous of Argentina, such as cranially positioned cervical neural spines and reduced metatarsal II, it is possible that the new Malagasy taxon represents a geographical extension of Noasauridae outside of Argentina. However, *Noasaurus*, as well as several other small-bodied, Late Cretaceous Gondwanan theropods from Argentina¹⁸ and India¹⁹, are represented by only a handful of elements, and thus the phylogenetic relationships of these taxa remain obscure.

The placement of *Masiakasaurus* in the Abelisauroidae indicates that small-bodied members of this clade, together with larger-bodied forms, spread throughout much of the southern supercontinent Gondwana by the Late Cretaceous. Thus, it seems that the Late Cretaceous radiation of North American coelurosaur theropods into small-bodied (ornithomimid, troodontid, dromaeosaurid) and large-bodied (tyrannosaurid) forms¹² occurred in parallel with the diversification of Gondwanan abelisauroid theropods. Furthermore, the peculiar dental and jaw morphology of *Masiakasaurus* suggests that we are still far from fully appreciating the morphological diversity of dinosaurs. □

Received 5 May; accepted 26 October 2000.

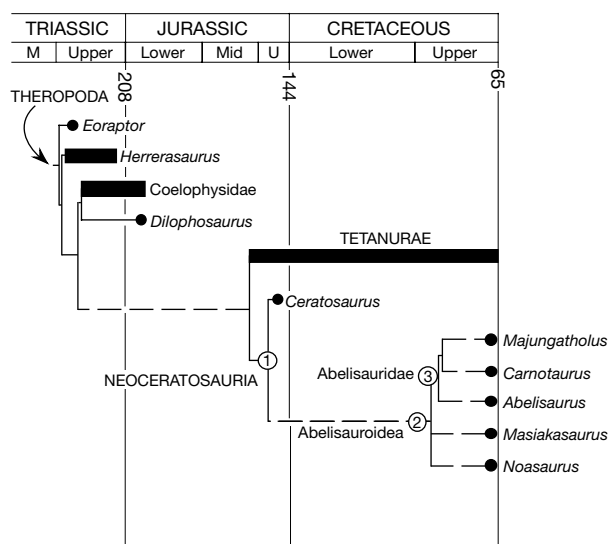


Figure 3 Stratigraphically calibrated cladogram of phylogenetic relationships of Abelisauroidae, including *Masiakasaurus*. Unambiguous synapomorphies supporting the nodes are: (1) (Neoceratosauria), lateral temporal fenestra much larger than orbit; numerous foramina surrounding axial diapophysis; notched posterior iliac margin; peg-and-socket articulation between ilium and pubis; femoral fourth trochanter reduced to low ridge; deeply excavated medial surface of proximal fibula; enlarged iliofibularis tubercle on fibular shaft; (2) (Abelisauroidae), caudal dentary margin with notch for surangular; enlarged external mandibular foramen; cervical neural spine abbreviated craniocaudally; seven sacral vertebrae; large rounded coracoid; humerus straight-shafted with bulbous head; hypertrophied mediodistal crest on femur; large tibial cnemial crest with laterally-directed hook; distal; cnemial crest curved dorsally; rectangular astragalar ascending process; triangular set of grooves on lateral and medial pedal unguals; and (3) (Abelisauridae), craniofacial elements with external sculpturing; long, shelf-like maxilla-jugal contact; greatly exposed paradental plates bearing a series of vertical ridges and grooves; dorsal vertebral parapophyses projecting far laterally from centrum; distal ends of caudal transverse processes expanded craniocaudally; forelimb short relative to other skeletal elements. M, Middle; U, Upper. Dates are millions of years before present.

- Bonaparte, J. F. & Powell, J. E. A continental assemblage of tetrapods from the Upper Cretaceous beds of El Brete, northwestern Argentina (Sauropoda-Coelurosauria-Carnosauria-Aves). *Mém. Soc. Géol. Fra.* **139**, 19–28 (1980).
- Bonaparte, J. F. The Gondwanian theropod families Abelisauridae and Noasauridae. *Hist. Biol.* **5**, 1–25 (1991).
- Bonaparte, J. F. Cretaceous tetrapods of Argentina. *Geowiss. Abhand.* **30**, 73–130 (1996).
- Krause, D. W., Prasad, G. V. R., von Koenigswald, W., Sahni, A. & Grine, F. E. Cosmopolitanism among Gondwanan Late Cretaceous mammals. *Nature* **390**, 504–507 (1997).
- Sampson, S. D. et al. Predatory dinosaur remains from Madagascar: implications for the Cretaceous biogeography of Gondwana. *Science* **280**, 1048–1051 (1998).
- Krause, D. W. et al. The Late Cretaceous vertebrate fauna of Madagascar: implications for Gondwanan paleobiogeography. *GSA Today* **9**, 1–7 (1999).
- Forster, C. A., Sampson, S. D., Chiappe, L. M. & Krause, D. W. The theropod ancestry of birds: new evidence from the Late Cretaceous of Madagascar. *Science* **279**, 1915–1919 (1998).
- Rogers, R. R., Hartman, J. H. & Krause, D. W. Stratigraphic analysis of Upper Cretaceous strata in the Mahajanga Basin, Madagascar: implications for ancient and modern faunas. *J. Geol.* **108**, 275–301 (2000).
- Sereno, P. C., Forster, C. A., Rogers, R. R. & Monetta, A. M. Primitive dinosaur skeleton from Argentina and the early evolution of Dinosauria. *Nature* **361**, 64–66 (1993).
- Gauthier, J. in *The Origin of Birds and the Evolution of Flight* (ed. Padian, K.) 1–47 (Memoirs of the California Academy of Sciences, San Francisco, 1986).
- Holtz, T. R. Jr The phylogenetic position of the Tyrannosauridae: implications for theropod systematics. *J. Paleontol.* **68**, 1100–1117 (1994).
- Sereno, P. C. The origin and evolution of dinosaurs. *Annu. Rev. Earth Planet. Sci.* **25**, 435–489 (1997).
- Sereno, P. C. The evolution of dinosaurs. *Science* **284**, 2137–2147 (1999).
- Rauhut, O. *Elaphrosaurus bambergi* and the early evolution of theropod dinosaurs. *J. Vert. Paleontol.* **18**, 71A (1998).
- Carrano, M. T. & Sampson, S. D. Evidence for a paraphyletic 'Ceratosauria' and its implications for theropod dinosaur evolution. *J. Vert. Paleontol.* **19**, 36A (1999).
- Forster, C. A. Gondwanan dinosaur evolution and biogeographic analysis. *J. Af. Earth Sci.* **28**, 169–185 (1999).
- Novas, F. E. in *The Encyclopedia of Dinosaurs* (eds Currie, P. J. & Padian, K.) 1–2 (Academic, New York, 1997).
- Bonaparte, J. F. Los vertebrados fósiles de la Formación Rio Colorado, de la Ciudad de Neuquén y cercanías, Cretácico Superior, Argentina. *Paleontología* **4**, 16–123 (1991).
- Huene, F. v. & Matley, C. A. The Cretaceous Saurischia and Ornithischia of the central provinces of India. *Palaeontol. Indica* **21**, 1–72 (1933).

Acknowledgements

We thank A. Rasoamimanana, B. Rakotosamimanana, P. Wright, B. Andriamihaja and the staff of the Institute for the Conservation of Tropical Environments for help with fieldwork in Madagascar; members of the 1993, 1995, 1996, 1998 and 1999 expeditions for their efforts; F. Novas and P. Currie for photographs of *Noasaurus*; D. Krause for comments on earlier drafts of the manuscript; V. Heisey for preparation; F. Grine for scanning electron microscope photography; and L. Betti-Nash and J. Higgins for assistance with the figures. This work was funded by grants from the National Science Foundation, the National Geographic Society, and the Dinosaur Society.

Correspondence and requests for materials should be addressed to S.D.S. (e-mail: ssampson@umnh.utah.edu).

Archaeal dominance in the mesopelagic zone of the Pacific Ocean

Markus B. Karner*, Edward F. DeLong† & David M. Karl*

* University of Hawai'i, Department of Oceanography, 1000 Pope Road, Honolulu, Hawai'i 96822, USA

† Monterey Bay Aquarium Research Institute, 7700 Sandholdt Road, Moss Landing, California 95039, USA

The ocean's interior is Earth's largest biome. Recently, cultivation-independent ribosomal RNA gene surveys have indicated a potential importance for archaea¹ in the subsurface ocean^{2–4}. But quantitative data on the abundance of specific microbial groups in the deep sea are lacking^{5,6}. Here we report a year-long study of the abundance of two specific archaeal groups (pelagic euryarchaeota and pelagic crenarchaeota)² in one of the ocean's largest habitats. Monthly sampling was conducted throughout the water column (surface to 4,750 m) at the Hawai'i Ocean Time-series

station⁷. Below the euphotic zone (> 150 m), pelagic crenarchaeota comprised a large fraction of total marine picoplankton, equivalent in cell numbers to bacteria at depths greater than 1,000 m. The fraction of crenarchaeota increased with depth, reaching 39% of total DNA-containing picoplankton detected. The average sum of archaea plus bacteria detected by rRNA-targeted fluorescent probes ranged from 63 to 90% of total cell numbers at all depths throughout our survey. The high proportion of cells containing significant amounts of rRNA suggests that most pelagic deep-sea microorganisms are metabolically active. Furthermore, our results suggest that the global oceans harbour approximately 1.3×10^{28} archaeal cells, and 3.1×10^{28} bacterial cells. Our data suggest that pelagic crenarchaeota represent one of the ocean's single most abundant cell types.

Sampling was carried out at roughly monthly intervals from September 1997 to December 1998 at the Hawai'i Ocean Time-series station ALOHA⁷ (22° 45' N, 158° 00' W) in the North Pacific subtropical gyre. We quantified cells belonging in the domain Bacteria, and two specific phylogenetic groups in the domain Archaea, which are prevalent in marine plankton: 'group 1 archaea' (referred to here as pelagic crenarchaeota); and 'group 2 archaea'

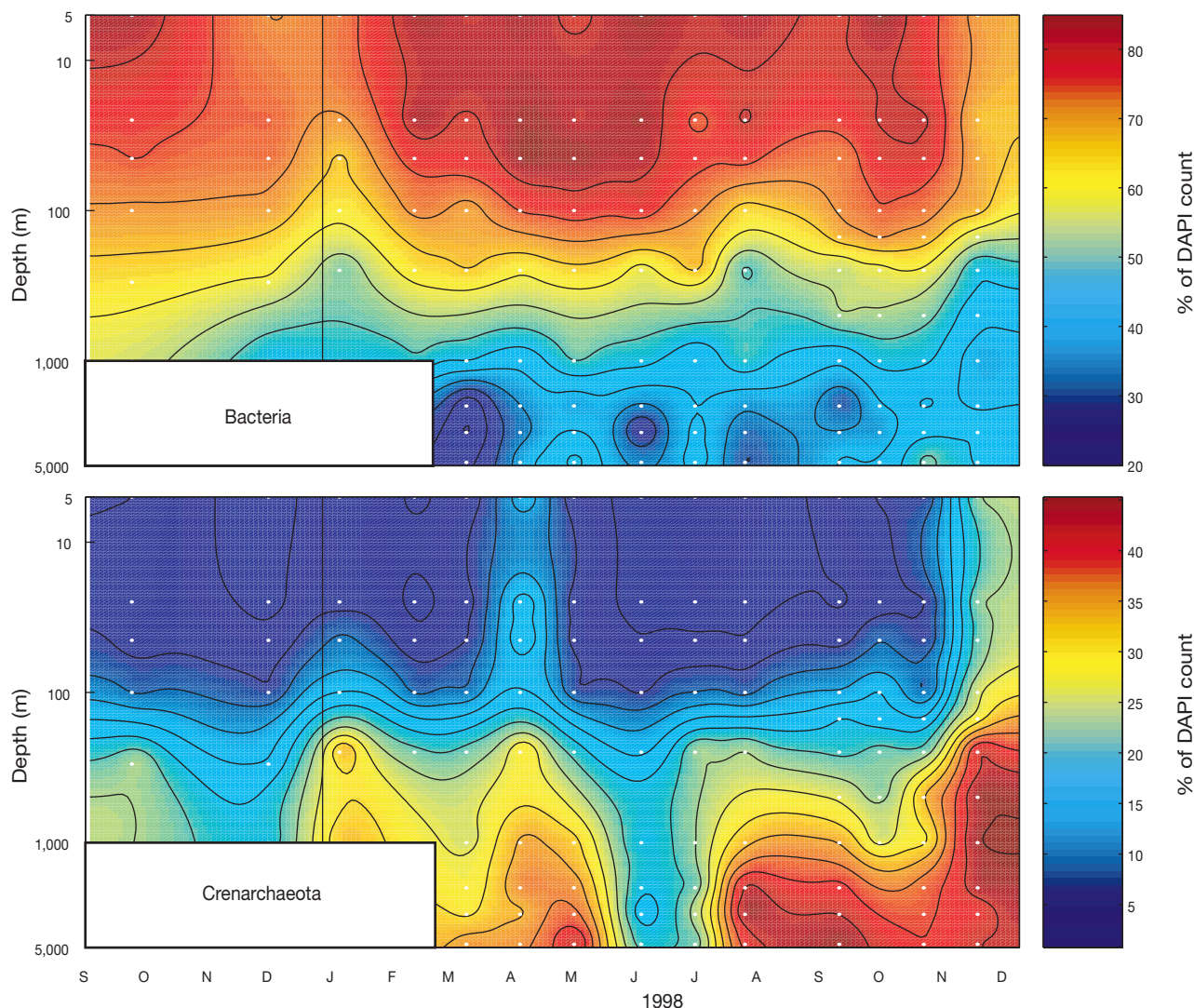


Figure 1 Contour plots of relative abundances with depth of bacteria and pelagic crenarchaeota during a 1-yr sampling effort at the Hawai'i Ocean Time-series station, ALOHA, in the North Pacific subtropical gyre. White dots indicate dates and depths where samples were collected. Contour lines are percentages of bacteria and pelagic

crenarchaeota as compared with total microbial abundance at each depth. Total cell abundance was assessed using the DAPI nucleic acid stain. Bacteria and archaea were enumerated using whole-cell rRNA targeted fluorescent *in situ* hybridization with fluorescein-labelled polynucleotide probes. See also Supplementary Information.

(referred to here as pelagic euryarchaeota)². Ribosomal RNA-targeted polynucleotide probes⁶ specific for either pelagic crenarchaeota, euryarchaeota or bacteria were hybridized with all samples collected from September 1997 until December 1998 (one deep sample collected in October 1998 had strong spurious background fluorescence and was omitted from the data set). Probe-binding cells were enumerated relative to total cells stained by 4',6-diamidino-2-phenylindole (DAPI)⁸ double staining⁹. Probe-conferred fluorescence was most intense in samples from shallower depths (0–500 m), and was weaker for both archaea and bacteria at depths greater than 500 m. Negative control counts (hybridization with a nonspecific probe and autofluorescence) were below 5% of total cells in 96% of all samples, hence a high degree of confidence in this method was achieved. (See also Supplementary Information.)

The two archaeal groups surveyed, as well as bacteria collectively, lacked clear seasonal trends in relative cellular abundance (Fig. 1). Therefore, averages were calculated for all samples for each depth layer, yielding the mean annual depth profile for each phylotype (Fig. 2a). Cells hybridizing with the bacterial probe dominated the population in the upper 150 m of the water column, representing up to 90% of all cells. Bacteria decreased in relative abundance with increasing depth (Figs 1, 2a), and below 1,000 m they represented only 35–40% of total cells. By comparison, pelagic crenarchaeota increased sharply in relative abundance at the 250-m depth layer, and below 1,000 m were as common as bacteria (Figs 1, 2a). Samples from 150 m (included in 4 of the 14 sampling dates) indicated that the relative increase in pelagic crenarchaeota occurred between 100 and 150 m, near the depth of the 1% isolume (Figs 1, 2a). In surface

layers, pelagic crenarchaeota were only present sporadically, and never abundant numerically. Pelagic euryarchaeota occasionally appeared in the near surface layer, but generally remained at a few per cent of the total count over the entire water column, too close to negative counts to establish a statistically reliable estimate (Fig. 2a). Pelagic crenarchaeota and euryarchaeota thus showed different patterns of abundance in the open sea (Fig. 2a), with low numbers of euryarchaeota in surface waters, in contrast to previous observations in coastal waters^{6,10,11}.

The sum of relative abundances of bacteria, pelagic euryarchaeota and pelagic crenarchaeota (subtracting negative controls) shows the total fraction of probe-binding cells over cells stained by DAPI (Fig. 2b). Although our deep-sea data on relative abundance of pelagic crenarchaeota (Fig. 2a) are globally higher than those recently reported in coastal waters⁶, our cumulative relative abundance data on total probe-positive cells (Fig. 2b) are lower. This may be due to the oligotrophic nature of the study site. Total cell abundance patterns reflected the higher total cell numbers in surface waters (Fig. 3). Total cell abundance declined by an order of magnitude between 150 and 1,000 m. Thus, the total number of pelagic crenarchaeotal cells peaked between 150 and 500 m, but remained at high relative numbers below this depth to the ocean floor.

In recent years cultivation-independent rRNA gene surveys have revealed new types of archaea in virtually every ecosystem examined^{12,13}. Yet marine planktonic archaeal types have eluded cultivation to date, and few quantitative data exist on their cellular distribution and abundance. Most previous studies have relied on the retrieval of rRNA-gene-containing clones to qualitatively describe naturally occurring picoplankton populations^{2–4}. A number of studies using quantitative rRNA hybridization techniques^{2,10,11,14,15} have suggested a widespread occurrence and high relative cell abundance of archaea in the upper water column. Unfortunately, these earlier studies do not allow for quantification of actual cell numbers. Quantitative microscopic surveys using single-cell, fluorescent *in situ* hybridization with rRNA-targeted oligonucleotide probes, by contrast, allow direct cell enumeration^{16,17}. A few investigations using oligonucleotide probes to enumerate marine picoplankton^{5,18–20} have succeeded in detecting

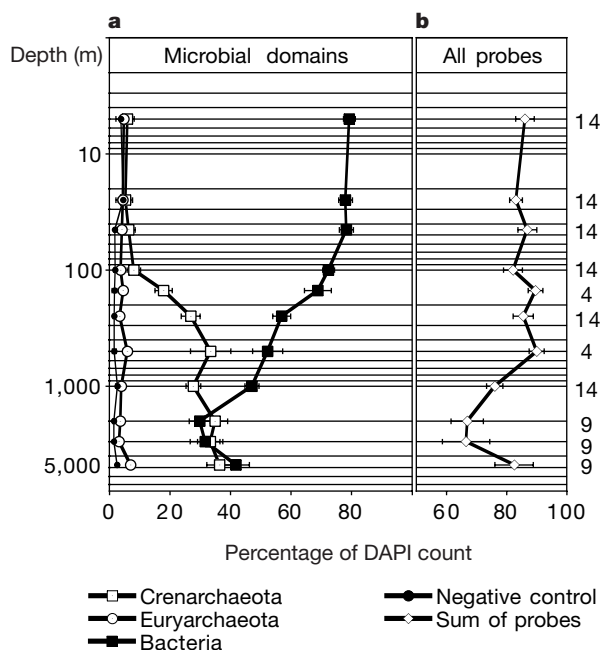


Figure 2 Mean annual depth profiles of microbial domains in the North Pacific subtropical gyre. Numbers are percentages of bacteria and archaea as compared to total microbial abundance at each depth. Total cell abundance was assessed using the DAPI nucleic acid stain. Bacteria and archaea were enumerated using whole-cell rRNA targeted fluorescent *in situ* hybridization with fluorescein-labelled polynucleotide probes. Data are averages of up to 14 roughly monthly samplings over a 1-yr period at the Hawai'i Ocean Time-series station, ALOHA. Error bars show standard error of mean; note column for total sample size at each depth. See also Supplementary Information. **a**, Depth profiles for bacteria (solid squares), pelagic crenarchaeota (open squares), pelagic euryarchaeota (open circles), and a non-specific control probe ('negative', solid circles). **b**, Depth profile of the sum of relative abundances of bacteria, pelagic crenarchaeota and pelagic euryarchaeota (open diamonds). Relative abundances of bacteria and both archaeal groups were summated at each depth and negative control data were subtracted.

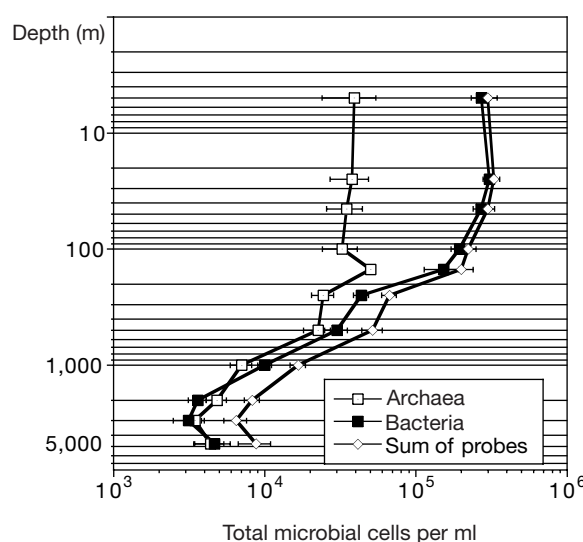


Figure 3 Mean annual depth profiles of microbial domains in the North Pacific subtropical gyre. Numbers are total cell abundances of bacteria and archaea (pelagic crenarchaeota and euryarchaeota combined). Bacteria and archaea were enumerated using whole-cell rRNA targeted fluorescent *in situ* hybridization with fluorescein-labelled polynucleotide probes. Data are averages of up to 14 roughly monthly samplings over a 1-yr period at the Hawai'i Ocean Time-series station, ALOHA. See also Supplementary Information.

naturally occurring bacterial and archaeal cells. One study⁵ using image intensification to enhance the fluorescence signal reported high percentages of Archaea (up to 60%) in coastal temperate waters and the Mediterranean Sea. Several studies, using an oligonucleotide probe targeting the domain Archaea, detected few or no archaeal cells in the upper water column^{19,20}.

None of these earlier studies, however, examined depths below 600 m, and only one¹⁸ identified specific archaeal cell types beyond the domain level. Further, single-cell hybridization studies have been hampered by the small size and apparent low rRNA content of marine picoplankton^{5,6,18}. Multiply labelled, rRNA-targeted polynucleotide probes⁶ allow more systematic enumeration of specific groups of planktonic microbes. But this approach has so far been demonstrated only at a coastal ocean site⁶, prompting our study at the Hawaii Ocean Time-series station.

In this report, we conclusively show that pelagic crenarchaeota were equivalent in numbers to bacteria throughout the entire meso- and bathypelagic zones in a major ocean gyre, over a 1-yr time-series survey. Two main points arise from our observations. First, on average, the fraction of pelagic crenarchaeota relative to DNA-containing microorganisms equals or exceeds the bacterial fraction below 1,000 m (Fig. 2a). This result suggests that pelagic crenarchaeota are a consistent and significant component of deep sea microbiota, and that they may rival bacterial abundances in the meso- and bathypelagic zones. Further, the standing stocks of bacteria and pelagic crenarchaeota were inversely proportional. Second, the total proportion of cells (bacterial and archaeal) detected with polynucleotide probes (Fig. 2b) remained fairly constant in surface waters to depths of 1,000 m, roughly 80% of the total DNA-containing cells⁸. At depths between 2,000 and 3,000 m, this number decreased to about 60% before increasing again to more than 70% close to the sea floor (Fig. 2b), similar to previously reported microbial cell numbers and metabolic activity profiles in the deep Pacific²¹. DNA-containing cells unlabelled by polynucleotide probes might include cells with low rRNA content, cells not recognized by the probes (for example, unusually small *Eucarya* in the 1- μ m range²²) or dead cells (ghosts)²³. Our results place an upper limit on the sum of the above categories. We conclude that the use of polynucleotide probes provides a robust approach for identification of deep-sea microbes, and that the fraction of remaining unlabelled cells remains below 20–30% of the total cells.

A previous study²⁴ suggested that the number of cells positive for rRNA hybridization techniques may reflect the number of metabolically active cells. If true, then the total number of probe-binding cells should represent the proportion of metabolizing cells in the water column (60–80% of total cells present). The decrease in probe-positive cells between 2,000 and 3,000 m might reflect decreasing availability or quality of growth substrate, resulting in a metabolically less active population. Close to the sea floor, gravitationally deposited organic matter might trigger the observed increase of the proportion of all probe-positive cells (Fig. 2b). This result is consistent with observed increases in metabolic activity near the sea floor previously found in Pacific bathypelagic microbial populations²¹. Our data indicate that most deep-sea microbes (archaea and bacteria) contain appreciable amounts of rRNA, and so may be active contributors to the ecosystem. Furthermore, given that the total number of microbes strongly decreases with depth in the ocean^{21,25}, our data lead to an ecologically significant conclusion: growth conditions for pelagic microbes do appear to be accurately reflected by microbial standing stock. There may also be a population shift with increasing depth to microbial groups better adapted to deep sea conditions, a large fraction of which appears to be pelagic crenarchaeota.

Combining our present data set with continental shelf data⁶, we can extrapolate total cell abundances for archaea to the global ocean²⁶. Such an extrapolation suggests that the world ocean contains approximately 1.3×10^{28} archaeal cells, and 3.1×10^{28}

bacterial cells. About 1.0×10^{28} cells, that is, 20% of all the picoplankton cells in the world ocean appear to be represented by one specific clade, the pelagic crenarchaeota. The habitat range for this single microbial group, spanning from mesopelagic to bathypelagic depths, is unusually broad. As a dominant component of the deep ocean, Earth's largest biome, archaea are thus far from confined to extreme niche habitats. Rather, the distribution of these archaea suggests that a common adaptive strategy has allowed them to radiate throughout nearly the entire oceanic water column. □

Methods

Sample preparation

Samples were fixed with 0.2- μ m filtered formalin (2% final concentration) and filtered onto 0.2- μ m polycarbonate filters. After filtration cells were treated by overlaying the filter for 2 min with a 0.25 M NaCl solution made up in 50% (v/v) ethanol. Samples were stored frozen either before or after filtration, with equal preservation of probing potential.

Whole-cell fluorescent *in situ* hybridization counts

Ribosomal RNA targeted hybridization was carried out on the filter-bound cells using fluorescein isothiocyanate (FITC) multiply labelled probes⁶ of more than 100 base pairs in length, targeting crenarchaeota (marine group 1 archaea) or euryarchaeota (marine group 2 archaea)¹⁰, or bacteria. A non-binding probe (consisting of the complementary strand of the pelagic crenarchaeotal probe) served as the negative control, while the summation of archaeal and bacterial fractions (minus negative control) provided a check on how total probe-positive cells compared with total cell counts using the DAPI stain⁸. Probe was added to a hybridization buffer at a final concentration of 2 ng μ l⁻¹. The hybridization buffer of 10% (w/v) dextran sulphate, 0.01% poly(A) and 0.1% sodium dodecyl sulphate (SDS) in 5 $\times$ SET was used with either 50% formamide (bacterial and negative control probes) or 70% formamide (for both archaeal probes, 1 $\times$ SET is 150 mM NaCl, 1 mM ethylenediaminetetraacetic acid (EDTA), 20 mM Tris, pH 8.0).

We did not automatically subtract the negative count corresponding to each sample because cells showing a 'negative' image, for example, owing to autofluorescence, could be of any group: crenarchaeota, euryarchaeota or bacteria. When reporting the sum of relative cell numbers from all probes, however, we did subtract the negative count as here we can relate the negative count to all cells, yielding a conservative estimate for total probe positive cells. During April 1998, high negative values of 25% and 36% were observed for the 5-m and the 25-m layers, respectively; for the other 120 samples the negative control averaged $2.0 \pm 1.5\%$ (mean $\pm$ s.d.). Hybridization temperatures were 65 °C for the two archaeal probes, and 55 °C for bacterial and negative control probes. Hybridization time was 12 h. Post-hybridization wash (2 h) was at 50 °C for archaeal probes, and 45 °C for bacterial/negative control probes in a 0.2 $\times$ SET solution made up in 50% formamide. Filters were mounted in Citifluor (Ted Pella), and FITC-positive cells were counted in relation to DAPI-positive cells on the same filter⁹. Counting was performed using a Zeiss epifluorescence microscope equipped with a $\times$ 100 Neofluar objective and 100-W mercury lamp illumination, as well as the appropriate filter sets for DAPI and FITC (Chroma Technology).

Received 4 July; accepted 21 November 2000.

- Woese, C. R., Kandler, O. & Wheelis, M. L. Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl Acad. Sci. USA* **87**, 4576–4579 (1990).
- DeLong, E. F. Archaea in coastal marine environments. *Proc. Natl Acad. Sci. USA* **89**, 5685–5689 (1992).
- Fuhrman, J. A., McCallum, K. & Davis, A. A. Novel major archaeobacterial group from marine plankton. *Nature* **356**, 148–149 (1992).
- Fuhrman, J. A. & Davis, A. A. Widespread Archaea and novel Bacteria from the deep sea as shown by 16S rRNA gene sequences. *Mar. Ecol. Prog. Ser.* **150**, 275–285 (1997).
- Fuhrman, J. A. & Ouverney, C. C. Marine microbial diversity studies via 16S rRNA sequences: cloning results from coastal waters and counting of native archaea with fluorescent single probes. *Aquat. Ecol.* **32**, 3–15 (1998).
- DeLong, E. F., Taylor, L. T., Marsh, T. L. & Preston, C. M. Visualization and enumeration of marine planktonic Archaea and Bacteria by using polyribonucleotide probes and fluorescent *in situ* hybridization. *Appl. Environ. Microbiol.* **65**, 5554–5563 (1999).
- Karl, D. M. & Lukas, R. The Hawaii Ocean Time-series (HOT) program: Background, rationale and field implementation. *Deep-Sea Res.* **43**, 129–156 (1996).
- Porter, K. & Feig, Y. S. The use of DAPI for identifying and counting aquatic microflora. *Limnol. Oceanogr.* **25**, 943–948 (1980).
- Hicks, R. E., Amann, R. I. & Stahl, D. A. Dual staining of natural bacterioplankton with 4', 6-diamidino-2-phenylindole and fluorescent oligonucleotide probes targeting kingdom-level 16S rRNA sequences. *Appl. Environ. Microbiol.* **58**, 2158–2163 (1992).
- Massana, R., Murray, A. E., Preston, C. M. & DeLong, E. F. Vertical distribution and phylogenetic characterization of marine planktonic Archaea in the Santa Barbara Channel. *Appl. Environ. Microbiol.* **63**, 50–56 (1997).
- Murray, A. E. et al. A time series assessment of planktonic archaeal variability in the Santa Barbara Channel. *Aquat. Microb. Ecol.* **20**, 129–145 (1999).
- Olsen, G. J. Archaea, archaea everywhere. *Nature* **371**, 657–658 (1994).
- Stein, J. L. & Simon, M. I. Archaeal ubiquity. *Proc. Natl Acad. Sci. USA* **93**, 6228–6230 (1996).
- DeLong, E., Wu, K. Y., Prézlin, B. B. & Jovine, R. V. M. High abundance of Archaea in Antarctic marine picoplankton. *Nature* **371**, 695–697 (1994).

15. Massana, R. *et al.* Vertical distribution and temporal variation of marine planktonic Archaea in the Gerlache strait, Antarctica, during early spring. *Limnol. Oceanogr.* **43**, 607–617 (1998).
16. Amann, R. L., Krumholz, L. & Stahl, D. A. Fluorescent-oligonucleotide probing of whole cells for determinative, phylogenetic, and environmental studies in microbiology. *J. Bacteriol.* **172**, 762–770 (1990).
17. Lee, S. & Kemp, P. F. Single-cell RNA content of natural marine planktonic bacteria measured by hybridization with multiple 16S rRNA-targeted fluorescent probes. *Limnol. Oceanogr.* **39**, 869–879 (1994).
18. Murray, A. E. *et al.* Seasonal and spatial variability of bacterial and archaeal assemblages in the coastal waters near Anvers island, Antarctica. *Appl. Environ. Microbiol.* **64**, 2585–2595 (1998).
19. Glöckner, F. O., Fuchs, B. M. & Amann, R. Bacterioplankton compositions of lakes and oceans: a first comparison based on fluorescence *in situ* hybridization. *Appl. Environ. Microbiol.* **65**, 3721–3726 (1999).
20. Simon, M., Glöckner, F. O. & Amann, R. Different community structure and temperature optima of heterotrophic picoplankton in various regions of the Southern Ocean. *Aquat. Microb. Ecol.* **18**, 275–284 (1999).
21. Nagata, T., Fukuda, H., Fukuda, R. & Koike, I. Bacterioplankton distribution and production in deep Pacific waters: Large-scale geographic variations and possible coupling with sinking particle fluxes. *Limnol. Oceanogr.* **45**, 426–435 (2000).
22. Courties, C. *et al.* Smallest eukaryotic organism. *Nature* **370**, 255 (1994).
23. Zweifel, U. L. & Hagström, Å. Total counts of marine bacteria include a large fraction of non-nucleoid-containing bacteria (ghosts). *Appl. Environ. Microbiol.* **61**, 2180–2185 (1995).
24. Karner, M. & Fuhrman, J. Determination of active marine bacterioplankton: a comparison of universal 16S rRNA probes, autoradiography, and nucleoid staining. *Appl. Environ. Microbiol.* **63**, 1208–1213 (1997).
25. Williams, P. M., Carlucci, A. F. & Olson, R. A deep profile of some biologically important properties in the central North Pacific gyre. *Oceanol. Acta* **3**, 471–476 (1980).
26. Menard, H. W. & Smith, S. M. Hypsometry of ocean basin provinces. *J. Geophys. Res.* **71**, 4305–4325 (1966).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank the crew of the RV *Moana Wave*, L. Tupas and K. Björkman for help with sampling. T. Taylor helped standardizing probing protocols, and L. Fujieki provided computational support. This study was supported by NOAA-Seagrant Office (MBK/DMK), NSF (DMK), and support from the David and Lucile Packard Foundation (EFD). This is SOEST contribution number 5313 and US JGOFS contribution number 647.

Correspondence and requests for materials should be addressed to M.B.K. (e-mail: mkarner@soest.hawaii.edu).

Testing Hamilton's rule with competition between relatives

Stuart A. West*, Martyn G. Murray*, Carlos A. Machado†, Ashleigh S. Griffin* & Edward Allen Herre‡

* Institute of Cell, Animal & Population Biology, University of Edinburgh, Edinburgh EH9 3JT, UK

† Department of Genetics, Rutgers University, Piscataway, New Jersey 08854, USA

‡ Smithsonian Tropical Research Institute, Apartado 2072, Balboa, Republic of Panama

Hamilton's^{1,2} theory of kin selection suggests that individuals should show less aggression, and more altruism, towards closer kin. Recent theoretical work has, however, suggested that competition between relatives can counteract kin selection for altruism^{3–11}. Unfortunately, factors that tend to increase the average relatedness of interacting individuals—such as limited dispersal—also tend to increase the amount of competition between relatives. Therefore, in most natural systems, the conflicting influences of increased competition and increased relatedness are confounded, limiting attempts to test theory^{4,8–10}. Fig wasp taxa exhibit varying levels of aggression among non-dispersing males that show a range of average relatedness levels. Thus, across species, the effects of relatedness and competition between relatives can be separated. Here we report that—contrary to Hamilton's original prediction^{1,2,12} but in agreement with recent

theory^{5–11}—the level of fighting between males shows no correlation with the estimated relatedness of interacting males, but is negatively correlated with future mating opportunities.

Hamilton's rule^{1,2} provides a tool for understanding a range of social interactions, including altruism, aggression, selfishness and spite. It states that altruism (or less aggression) is favoured when $rb - c > 0$, where c is the fitness cost to the altruist, b is the fitness benefit to the beneficiary and r is their genetic relatedness. For a given benefit and cost, the evolution of altruism therefore relies upon a sufficiently high relatedness between interacting individuals. Hamilton² originally suggested that a high relatedness could arise in two ways: (1) behaviour based upon direct kin recognition between individuals, or (2) limited dispersal (population viscosity).

However, the importance of limited dispersal in increasing the relatedness among interacting individuals and favouring altruism has been controversial^{3–11}. Hamilton's original suggestion has been contested because limited dispersal can also increase competition between neighbouring relatives, which opposes the evolution of altruistic behaviour^{3–11}. Unfortunately, empirical tests of theory, that determine the relative importance of increases in both relatedness and competition between relatives, have been hindered because both factors are influenced by dispersal, and so their effects are usually confounded^{14,8–10}.

The variable form of mate competition and population structure across fig wasp species with wingless males offers an opportunity for disentangling the confounded effects of relatedness and competition between relatives in viscous populations^{12–18}. Fig wasps are species that develop within the fruit of fig trees, and include mutualistic pollinating species as well as parasitic non-pollinating species¹⁵. In many species the males are wingless, and mate with the winged females before the females disperse. The level of aggression between these non-dispersing males varies enormously across species^{12–15,18}. At one extreme, males of some non-pollinating species are highly modified for combat with armoured bodies and huge mandibles. These mandibles are used to tear soft tissue and sever body parts, including limbs, head and abdomen, and can result in extremely high mortality levels. At the other extreme, males of other non-pollinating and most pollinating species show no modifications for combat or aggression.

Across these species, the average relatedness of competing males varies enormously owing to variation in the number of females that lay eggs in each fruit^{12–16}. For example, if only one female lays eggs in a fruit then all the competing males will be brothers; increasing numbers of females laying eggs in a fruit will lead to males

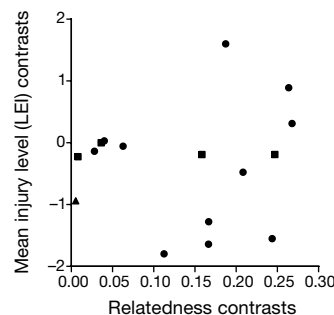


Figure 1 Mean injury level contrasts plotted against estimated relatedness contrasts. Across species, the mean injury level (lifetime extent of injury, LEI) and the proportion of individuals severely injured (SI) showed no significant relationship with estimated relatedness (LEI: all contrasts, $F(1,15) = 1.01$, $r^2 = 0.06$, $p = 0.33$; not including contrasts within the pollinator lineage, $F(1,11) = 0.72$, $r^2 = 0.06$, $p = 0.42$; SI: all contrasts, $F(1,15) = 0.04$, $r^2 < 0.01$, $p = 0.84$; not including contrasts within the pollinator lineage, $F(1,11) = 0.05$, $r^2 < 0.01$, $p = 0.83$). Circles, contrasts between the non-pollinating species; squares, contrasts between the pollinator species; triangle, the contrast between pollinators and non-pollinators.

competing with unrelated individuals^{12–16}. Hamilton¹² suggested that the variation in the level of fighting across fig wasp species (and other insects with fighting males) reflects different levels of relatedness: specifically, that kin selection favours altruism and less aggression in species in which competing males are more highly related. However, this prediction did not take into account the potentially conflicting effect of competition between relatives.

Wingless male fig wasps represent an extreme case with respect to competition between relatives because there is no dispersal before competition (competition is completely local)^{9–11}. More recent theory suggests that in this case competition between relatives totally removes the kin-selected benefits of altruism towards relatives^{5–11}, and so increased relatedness will not favour lower levels of aggression. Instead we predict that fighting levels should correlate with the importance of fighting over (and therefore mating with) any particular female^{13,17}, and so be negatively correlated to the total number of females developing in a fruit.

We estimated fighting levels and average relatedness between competing males in 25 fig wasp species (see Methods). Fighting levels were estimated by scoring wingless males for injuries after females had left the fruit¹³. Individual injuries were rated on a scale of 0–8 (for example, loss of an antenna scored 0.5 points, whereas loss of head with evisceration scored 8.0 points)¹³. The overall lifetime extent of injury (LEI) score for each individual was found by adding together the scores from its different injuries, subject to a maximum of 8.0. Individuals that had injuries rated at 8 points were classed as severely injured (SI). We used the sex ratio (proportion of males) to estimate average relatedness between competing males for the different wasp species (see Methods). Abundant evidence in pollinating and non-pollinating fig wasps suggests that the sex ratio of a species correlates with the number of females that lay eggs per fruit, and therefore the relatedness of individuals within fruits^{15,16}.

Previous comparative studies of fighting levels between males across fig wasp species^{12,14,18} have been hindered by the problem that male morphology, levels of fighting and the possible correlates of fighting levels (relatedness, male and female density) are tightly linked to phylogeny^{14–16,18}. Pollinating species (subfamily Agaoninae) generally develop in large broods, and competing males are often highly related and do not fight violently. In contrast, non-pollinating species (subfamilies Epichrysomallinae, Otitesellinae, Sycoecinae, Sycoryctinae and Sycophaginae) frequently develop in smaller broods, competing males can be less related and, in several cases, exhibit violent fighting. This is a prime model for when

phylogenetic differences may lead to misleading correlations¹⁹; the potential danger of this is clearly shown by the fact that subfamily is able to explain greater than 50% of the variation in fighting levels across species. We address this potential problem by using a formal comparative method (independent contrasts¹⁹) based upon a molecular phylogeny that we have estimated (see Methods).

Controlling for phylogeny we found that both the mean LEI and the proportion of SI individuals showed no significant relationship with estimated relatedness (Fig. 1). In contrast, both LEI and the proportion of SI individuals were significantly negatively correlated with the mean number of females developing in a fruit (Fig. 2), as predicted by contest models which do not include any effect of relatedness^{13,17}. Thus, while the sex ratio of the wasps is correlated to relatedness and possible outcrossing^{10,12,15,16}, the form of mate competition between males responds to a different aspect of population structure—the number of females developing in a fruit. The robustness of our results is supported by the fact that the same results were also observed when the data were analysed without the contrasts derived within the pollinator lineage (where no fighting has been observed in the species we considered; Fig. 2). Furthermore, when testing for an effect of relatedness, our comparison across species provides more power than a comparison within a species (across figs), which would also require individuals to be able to assess their relatedness to other males in the fruit (some form of kin recognition) and adjust their behaviour accordingly.

Our results support the prediction that, with limited dispersal, the increased competition between relatives can negate the effect of increased relatedness in favouring altruism^{5–11}. One way¹⁰ of incorporating this into Hamilton's rule ($rb - c > 0$) is by expressing the marginal benefit of increased altruism, b , as a function of three parameters: $b = B - a(B - c)$ (Fig. 3; for alternative methods which focus on how relatedness is measured, see ref. 9). Here c has the standard meaning of the cost of altruism to the actor, and B is the benefit that would accrue to the recipients if the recipients did not compete with each other. The parameter a is the spatial scale at which competition occurs: an increase in the reproductive success of neighbours by a proportion x increases local competition by a factor ax , but has negligible effect on the intensity of global competition because the local neighbourhood is only a small part of the total population¹⁰. The parameter a therefore measures the extent to which neighbours (and potentially relatives) compete. If competition is completely global ($a = 0$) then any competition between relatives is negligible, and so the classic equation for Hamilton's rule

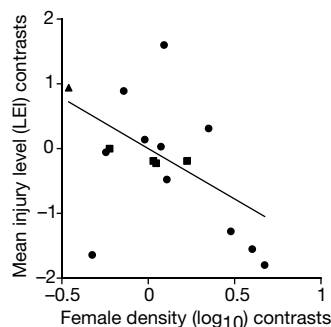


Figure 2 Mean injury level (LEI) contrasts plotted against the mean number of females developing in a fruit ($\log_{10}$ transformed) contrasts. Across species, the mean injury level (LEI) and the proportion of individuals severely injured (SI) were significantly negatively correlated with the mean number of females developing in a fruit ($\log_{10}$ transformed for the analyses containing all contrasts; LEI: all contrasts, $F(1,15) = 5.92$, $r^2 = 0.28$, $P = 0.03$; not including contrasts within the pollinator lineage, $F(1,11) = 6.84$, $r^2 = 0.38$, $P = 0.02$; SI: all contrasts, $F(1,15) = 3.01$, $r^2 = 0.17$, $P = 0.10$; not including contrasts within the pollinator lineage, $F(1,11) = 5.00$, $r^2 = 0.31$, $P = 0.047$). See Fig. 1 for symbol (circles, squares and triangle) definitions.

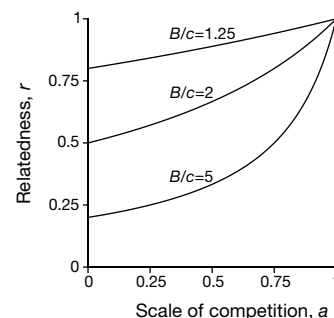


Figure 3 Hamilton's rule with relatedness and competition between relatives. Plotted is the minimum relatedness (r) required to favour altruism, against the scale of competition (a) (ref. 10). Competition varies from completely global ($a = 0$) to completely local ($a = 1$). The different lines represent different ratios of B/c , where B is the fitness benefits to the beneficiaries in the absence of competition between relatives, and c is the cost of altruism to the actor. We note that: (1) as competition becomes more local (greater competition between relatives; higher a), the relatedness (r) or relative fitness benefit to the beneficiary (B/c) must increase in order to favour altruism; (2) when competition is completely local ($a = 1$) altruism cannot spread ($r > 1$ is required).

holds^{9,10}. As competition becomes more local (increasing a), any increase in reproduction of a neighbour comes at a greater cost to other neighbours. This increases competition between relatives and so reduces the kin selection advantage in being altruistic (reduces b). In the extreme, with competition being completely local ($a = 1$), as is the case with competition for mates between wingless male fig wasps, the increased competition between relatives exactly cancels the advantage of increased relatedness. At this point any increase in the reproduction of a neighbour must come at the cost of other related neighbours, and so altruism is not favoured^{9,10}. Further support for this prediction comes from the observation that, in other insect cases in which closely related males show extreme aggression, competition for mates tends to be very local ($a \rightarrow 1$)¹².

More generally, our results emphasize the importance of determining the amount of competition between relatives, and that estimating the scale at which competition occurs (a) can be as crucial as estimating relatedness (r). Whereas the importance of relatedness is well appreciated, and frequently estimated, the scale at which competition occurs (a) is generally ignored, let alone quantitatively estimated. Although this may partly reflect difficulties in empirical estimation of a , its effect may be incorporated into Hamilton's rule in several ways^{9,10}. When competition between relatives is ignored, the importance of kin selection is overestimated. Furthermore, there are a number of recent studies that fail to show a correlation between relatedness and altruism²⁰, including behaviours such as cannibalism (insects²¹), cooperative breeding (meerkats and several bird species^{20,22,23}) and watching for predators (meerkats²⁴). We suggest that some of these cases might be explained by the fact that there is enough competition between relatives to ensure that kin selection benefits are relatively unimportant (sufficiently high a). Consistent with this possibility is that, in all of these cases, an individual (direct) fitness explanation can be provided for the behaviour in question and its variation. □

Methods

Data collection

Data on 25 species of fig wasps had been collected from seven species of Malaysian wild fig^{13,14}. The wasp species consist of eight pollinator species, and 17 non-pollinator species. Almost-ripe fruit were collected, from which wasps were about to emerge. The fruit were enclosed in jars with cloth lids, and after the wasps emerged, the contents of each jar were preserved in 75% alcohol. Later, all the wasps which had and had not emerged from the fruit were identified to species, sexed, counted and the males scored for the injuries that they obtained in their lifetime (termed lifetime extent of injury, LEI). Individual injuries were rated on a scale of 0–8 (loss of part or whole antenna, 0.5 points; loss of part or whole tarsus or small bruise, 1.0; loss of part or whole tibia (plus tarsus) or bruise with cut, 2.0; loss of part or whole femur (plus tibia and tarsus) or large bruise or bruise plus crushed area, 3.0; loss of part or whole coxa (plus femur, tibia and tarsus) or half severed abdomen or head with no evisceration, 4.0; >half severed abdomen or head with evisceration, 8.0)¹³.

We used the sex ratio (proportion of males) to estimate average relatedness between competing males for the different wasp species. In both pollinating and non-pollinating fig wasps, the sex ratio varies with the number of females that laid eggs in a fruit, and therefore average relatedness, as predicted by sex ratio theory^{12,15,16}. The fewer females that lay eggs in a fruit, the more female biased their offspring sex ratio is. Quantitative estimates of mean relatedness between wasps within a fruit can be obtained by rearranging a well-known expression for the expected (unbeatable) sex ratio (the proportion of males, m) as a function of relatedness (r) in haplodiploid species^{10,12,15,16}, to give r as a function of m : $r = 4(1 - m)/(3 - m + \sqrt{1 + 10m + m^2})$. More generally, the use of sex ratio data to estimate relatedness has gained quantitative support from work on other organisms (malaria and related protozoan parasites), where estimates of relatedness derived from sex ratios have been confirmed by estimates derived directly from genotypes^{25,26}.

Phylogeny

A molecular phylogeny was estimated using nucleotide sequences 816 base pairs long collected from the 3' end of the cytochrome oxidase subunit I gene (COI)^{27,28}. Analyses were performed with version 4.0b1 of PAUP*, written by D. L. Swofford (Smithsonian Institution, Washington, DC). Our molecular data set contained 15 species of fig wasps that included either the species scored for male injury level, or species from the same genera or subfamily (GenBank accession numbers: AF302052–AF302067). Sequences from the following species were used in the phylogenetic analyses: *Ceratosolen constrictus*, *Ceratosolen solmsi* (same species scored for male injury level); *Blastophaga nipponica*, *Eupristina verticillata*, *Ceratosolen galili*, *Eujacobsonia* sp. (light trap, unknown host), *Apocrypta* sp. (ex. *Ficus sycamorius*), *Sycorictes* sp. (ex. *Ficus glumosa*), *Sycoscapter* sp.

(ex. *F. cerecarpa*), *Philotrypes* sp. (ex. *Ficus tinctoria*), *Eukoebelia* sp. (ex. *F. sycamorius*) (species of the same genera as those scored for injury level); *Heterandrium* sp. (ex. *Ficus dugandii*) (non-pollinator from subfamily Otitesellinae). Additionally, sequences from one species of the African genus *Apocryptophagus* (ex. *F. sycamorius*) and two species from the neotropical genus *Idarnes* (ex. *Ficus trigonata*, *Ficus obtusifolia*) were used in the phylogenetic analyses to balance the uneven sampling from different subfamilies. Maximum likelihood methods were used to reconstruct the phylogeny. The general reversible model with rate heterogeneity was used²⁹, and the parameters of the model were estimated from the data. The tree topology was estimated using an heuristic algorithm with branch swapping (tree bisection–reconnection). The heuristic search was repeated three times and the consensus tree from the best three trees was used for generating the final phylogeny (see Supplementary Information). The species scored for male injury level that were not present in the molecular phylogeny were added to the nodes leading to species of the same genus or subfamily.

Statistical analyses

The data were analysed using phylogenetically independent contrasts. These were derived by calculating the difference in the response and explanatory variables across pairs of species or higher nodes that share a common ancestor¹⁹, as implemented in the CAIC statistical package³⁰. The continuous explanatory variables used were those that theory suggests may be important: the average number of males in a fruit, the average number of females in a fruit, the ratio of average number of females to average number of males, 1/(the average number of females times the number of males), and estimated relatedness. The explanatory powers of these different variables were tested without transformation, log₁₀ transformed, and inverted. Our phylogeny led to 16 contrasts from the 25 species. These contrasts were then analysed with multiple regressions through the origin¹⁹. In all analyses we initially fitted a full model, including all explanatory variables that we were considering. Terms were then removed from the model by stepwise deletion. Across species, the absolute (untransformed) ranges for the various variables were: number of females developing in a fruit (1.7–574.5); LEI (0.046–5.496); SI (<0.01–0.48). The mean number of females developing in a fruit showed no significant relationship with the mean sex ratio ($F(1,23) = 1.31$, $r^2 = 0.05$, $p = 0.26$).

Received 20 June; accepted 13 November 2000.

- Hamilton, W. D. The evolution of altruistic behaviour. *Am. Nat.* **97**, 354–356 (1963).
- Hamilton, W. D. The genetical evolution of social behaviour, I & II. *J. Theor. Biol.* **7**, 1–52 (1964).
- Grafen, A. in *Behavioural Ecology: An Evolutionary Approach* (eds Krebs, J. R. & Davies, N. B.) 62–84 (Blackwell Scientific, Oxford, 1984).
- Murray, M. G. & Gerrard, R. J. Conflict in the neighbourhood: models where close relatives are in direct competition. *J. Theor. Biol.* **111**, 237–246 (1984).
- Wilson, D. S., Pollock, G. B. & Dugatkin, L. A. Can altruism evolve in purely viscous populations. *Evol. Ecol.* **6**, 331–341 (1992).
- Taylor, P. D. Altruism in viscous populations—an inclusive fitness model. *Evol. Ecol.* **6**, 352–356 (1992).
- Taylor, P. D. Inclusive fitness in a homogeneous environment. *Proc. R. Soc. Lond. B* **249**, 299–302 (1992).
- Queller, D. C. Does population viscosity promote kin selection? *Trends Ecol. Evol.* **7**, 322–324 (1992).
- Queller, D. C. Genetic relatedness in viscous populations. *Evol. Ecol.* **8**, 70–73 (1994).
- Frank, S. A. *Foundations of Social Evolution* (Princeton Univ. Press, Princeton, 1998).
- Kelly, J. K. The effect of scale dependent processes on kin selection: mating and density regulation. *Theor. Popul. Biol.* **46**, 32–57 (1994).
- Hamilton, W. D. in *Reproductive Competition and Sexual Selection in Insects* (eds Blum, M. S. & Blum, N. A.) 167–220 (Academic, New York, 1979).
- Murray, M. G. The closed environment of the fig receptacle and its influence on male conflict in the Old World fig wasp, *Philotrypes pilosa*. *Anim. Behav.* **35**, 488–506 (1987).
- Murray, M. G. Environmental constraints on fighting in flightless male fig wasps. *Anim. Behav.* **38**, 186–193 (1989).
- Herre, E. A. et al. in *Social Competition and Cooperation in Insects and Arachnids Vol. I, The Evolution of Mating Systems* (eds Choe, J. & Crespi, B.) 226–239 (Princeton Univ. Press, Princeton, 1997).
- West, S. A. & Herre, E. A. Partial local mate competition and the sex ratio: a study on non-pollinating fig wasps. *J. Evol. Biol.* **11**, 531–548 (1998).
- Enquist, M. & Leimar, O. The evolution of fatal fighting. *Anim. Behav.* **39**, 1–9 (1990).
- Vincent, S. L. *Polymorphism and Fighting in Male Fig Wasps*. Thesis, Rhodes Univ. (1991).
- Harvey, P. H. & Pagel, M. D. *The Comparative Method in Evolutionary Biology* (Oxford Univ. Press, Oxford, 1991).
- CluttonBrock, T. H. et al. Individual contributions to babysitting in a cooperative mongoose, *Suricata suricatta*. *Proc. R. Soc. Lond. B* **267**, 301–305 (2000).
- Carcamo, H. A. & Spence, J. R. Kin discrimination and cannibalism in water striders (Heteroptera: Gerridae): another look. *Oikos* **70**, 412–416 (1994).
- Dunn, P. O., Cockburn, A. & Mulder, R. A. Fairy-wren helpers often care for young to which they are unrelated. *Proc. R. Soc. Lond. B* **259**, 339–343 (1995).
- Magrath, R. D. & Whittingham, L. A. Subordinate males are more likely to help if unrelated to the breeding female in cooperatively breeding white-browed scrubwrens. *Behav. Ecol. Sociobiol.* **41**, 185–192 (1997).
- CluttonBrock, T. H. et al. Selfish sentinels in cooperative mammals. *Science* **284**, 1640–1644 (1999).
- Read, A. F., Anwar, M., Shutler, D. & Nee, S. Sex allocation and population structure in malaria and related parasitic protozoa. *Proc. R. Soc. Lond. B* **260**, 359–363 (1995).
- West, S. A., Smith, T. G. & Read, A. F. Sex allocation and population structure in apicomplexan (protozoa) parasites. *Proc. R. Soc. Lond. B* **267**, 257–263 (2000).
- Machado, C. A., Herre, E. A., McCafferty, S. & Bermingham, E. Molecular phylogenies of fig pollinating and non-pollinating wasps and the implications for the origin and evolution of the fig–fig wasp mutualism. *J. Biogeogr.* **23**, 531–542 (1996).
- Machado, C. A. *Molecular Natural History of Fig Wasps*. Thesis, Univ. California (1998).
- Yang, Z. Estimating the pattern of nucleotide substitution. *J. Mol. Evol.* **39**, 105–111 (1994).

30. Purvis, A. & Rambaut, A. Comparative analysis by independent contrasts (CAIC): an Apple Macintosh application for analysing comparative data. *Comput. Applics Biosci.* **11**, 247–251 (1995).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank S. Frank, S. Nee, S. Reece and R. Trivers for comments that improved the clarity of our manuscript. This work was supported by the BBSRC, STRI and NERC.

Correspondence and requests for materials should be addressed to S.A.W. (e-mail: stu.west@ed.ac.uk).

Reduced antinociception and plasma extravasation in mice lacking a neuropeptide Y receptor

Philippe Naveilhan^{*†}, Heshameh Hassani^{*†}, Guilherme Lucas^{*}, Karin Hygge Blakeman[‡], Jing-Xia Hao[‡], Xiao-Jun Xu[‡], Zsuzsanna Wiesenfeld-Hallin[‡], Peter Thorén[§] & Patrik Ernfors^{*}

^{*} Laboratory of Molecular Neurobiology, Department of Medical Biochemistry and Biophysics, and [§] Department of Physiology and Pharmacology, Karolinska Institute, S-17177 Stockholm, Sweden

[‡] Department of Medical Laboratory Sciences and Technology, Division of Clinical Neurophysiology, Huddinge University Hospital, Karolinska Institute, S-14186 Huddinge, Sweden

[†] These authors contributed equally to this work

Neuropeptide Y (NPY) is believed to exert antinociceptive actions by inhibiting the release of substance P and other 'pain neurotransmitters' in the spinal cord dorsal horn^{1–3}. However, the physiological significance and potential therapeutic value of NPY remain obscure⁴. It is also unclear which receptor subtype(s)

are involved. To identify a possible physiological role for the NPY Y1 receptor in pain transmission, we generated NPY Y1 receptor null mutant (Y1^{−/−}) mice by homologous recombination techniques. Here we show that Y1^{−/−} mice develop hyperalgesia to acute thermal, cutaneous and visceral chemical pain, and exhibit mechanical hypersensitivity. Neuropathic pain is increased, and the mice show a complete absence of the pharmacological analgesic effects of NPY. In the periphery, Y1 receptor activation is sufficient and required for substance P release and the subsequent development of neurogenic inflammation and plasma leakage. We conclude that the Y1 receptor is required for central physiological and pharmacological NPY-induced analgesia and that its activation is both sufficient and required for the release of substance P and initiation of neurogenic inflammation.

Homologous recombination in embryonic stem cells was used to establish mice deficient in the NPY Y1 receptor. The disruption was generated by introducing an internal ribosomal entry site followed by a Tau–LacZ fusion minigene into the second exon of Y1 (Fig. 1a). Southern blot analysis confirmed that the Y1 allele was disrupted and northern blot analysis showed that instead of the messenger RNA transcripts encoding Y1, the mutant (Y1^{−/−}) mice produced the expected mRNA encoding β -galactosidase (Fig. 1b–d). As previously described, female Y1^{−/−} mice display a late-onset overweight compared to their littermates⁵ (data not shown). Y1 receptors are abundant in the forebrain, whereas little or nothing is present in the brainstem⁶. Y1 receptors are also highly expressed in dorsal root ganglion neurons in preferentially small and medium size neurons^{6,7}. However, the central termination of Y1 nerve fibres in the dorsal horn, and whether Y1 is expressed in both of the two major cytochemical subpopulations of pain neurons, the SP peptidergic and non-peptidergic pain neurons⁸, is unresolved.

β -galactosidase histochemical and immunohistochemical staining of spinal cord sections from Y1^{−/−} mice showed strong staining localized exclusively to the dorsal horn (Fig. 1e, f). Immunohistochemical double staining for β -galactosidase (staining Y1 expressing neurons and fibres) and the lectin IB4 (staining somas and nerve fibres of unmyelinated non-peptidergic sensory nociception neurons⁸) showed a strong staining for Y1 nerve fibres in dorsal

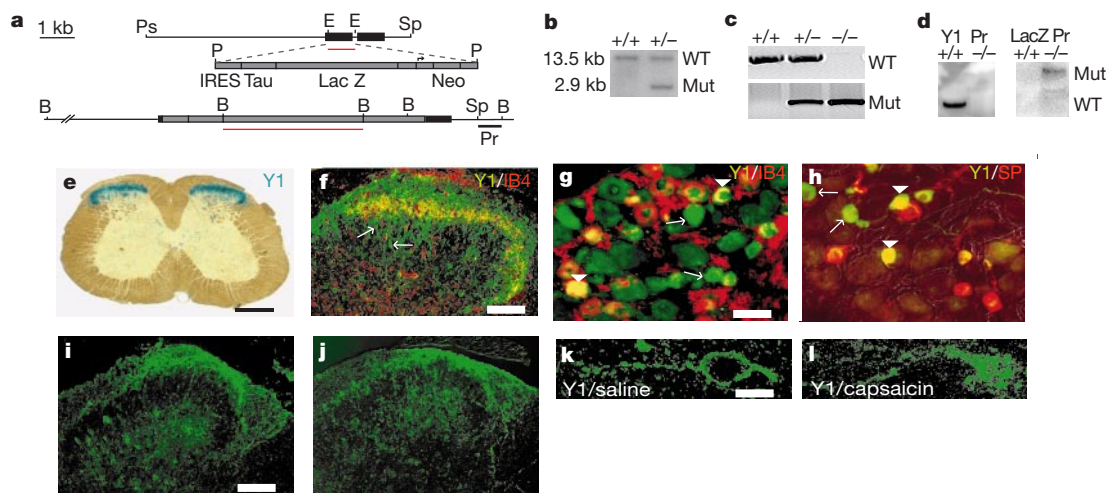


Figure 1 Targeted mutagenesis of the Y1 receptor and expression analysis of Y1 and SP receptors. **a**, Y1 gene targeting. Top, targeting vector (black boxes, Y1 coding exons). The disrupting cassette is indicated. Bottom, restriction map of the resulting targeted allele (B, *Bam*HI; Sp, *Spe*I; E, *Eco*RI; P, *Pac*I; Pr, probe used in the Southern blots.). **b**, Southern blot of ES cells. **c**, PCR genotyping of wild-type, Y1^{+/−} and Y1^{−/−} mice. **d**, Northern blot of total brain RNA of Y1^{+/−} and Y1^{−/−} mice using a Y1 probe (Y1 Pr) or LacZ probe (LacZ Pr). These probes are underlined in red in **a**. **e**, Transverse section from the spinal cord lumbar enlargement of Y1^{+/−} mice histochemically stained for β -galactosidase. **f**, Y1^{+/−} mice immunohistochemically stained for β -galactosidase-positive nerve terminals and

neurons (arrows) in the spinal cord dorsal horn (green) and the lectin IB4 (red, layer I liner). **g**, **h**, Double staining of L4 dorsal root ganglion for β -galactosidase (green) and IB4 (red) (**g**), and for β -galactosidase (green) and SP (red) (**h**). Arrows point to single-stained neurons, arrowheads to double-stained neurons. **i**, **j**, SP receptor distribution in the dorsal horn of wild-type mice (**i**) and Y1^{−/−} mice (**j**). **k**, SP receptor staining in lamina I of the contralateral vehicle injected side of Y1^{−/−} mice. **l**, Loss of cell surface and increase of intracellular SP receptor immunoreactivity in lamina I 10 min after capsaicin injection into the hindpaw of Y1^{−/−} mice. Scale bar, 300 μ m (**e**), 80 μ m (**f**, **i**, **j**), 30 μ m (**g**, **h**), 20 μ m (**k**, **l**).

horn layer II overlapping with IB4 terminals. A significant portion of the nerve fibres also terminated in layers I, III and IV (Fig. 1f). Y1-positive dorsal horn interneurons were also found (Fig. 1f, arrows). Many Y1-expressing dorsal root ganglion neurons co-expressed substance P (SP) (Fig. 1h; 38% of Y1 neurons contained SP), however, a large number of Y1 neurons also double stained for IB4 (Fig. 1g; 32%).

A presynaptic block of primary afferent SP release in the spinal cord may participate in central NPY-induced analgesia³. We therefore examined whether this primary afferent circuit was intact in the Y1^{-/-} mice. No overt alteration of SP nor NPY immunoreactivity was detected in the spinal cord of Y1^{-/-} mice (data not shown). Furthermore, total SP in spinal cord measured by enzyme immunoassay (EIA) was $2,852 \pm 328.9 \text{ pg g}^{-1}$ tissue and $3,919 \pm 444.1 \text{ pg g}^{-1}$ tissue in wild-type and Y1^{-/-} mice, respectively ($P > 0.05$, Student's *t*-test). Immunohistochemical detection of the SP receptor revealed similar staining in the somatic and dendritic cell surface of neurons in the superficial spinal cord lamina (I and II) of both wild-type and Y1^{-/-} mice, as in previous results⁹ (Fig. 1i, j). Substance P release leads to SP receptor activation and internalization¹⁰. Neurons containing internalized SP receptors were detected in Y1^{-/-} mice

after an intraplantar injection of capsaicin (Fig. 1k, l), indicating that there is no defect of receptor activation *per se* in the absence of the Y1 receptor. Thus, we conclude that the main neuronal pain circuit suggested to be modulated by NPY is anatomically intact in the Y1^{-/-} mice.

Behavioural acute nociceptive thresholds were affected markedly by an absence of Y1. Substance P/neurokinin A null mice show a blunted response to painful stimulus only at moderate intensities, whereas response to mildly or intensely painful stimulus is intact¹¹. We therefore tested whether also the NPY Y1 receptor functions at specific thresholds in modulating nociception. Withdrawal latency in the hot-plate assay was examined at 48, 50, 52, 55 and 58 °C. At 48 °C, no significant difference was observed (data not shown). The Y1^{-/-} mice showed, however, a profound hyperalgesia and displayed a significantly reduced latency at all temperatures above 48 °C (Fig. 2a). The hot-plate test involves supraspinal integration associated with the paw withdrawal. To test whether neuropeptide Y1 receptor might be acting in a spinal circuitry, we characterized these mice in the tail-flick assay. Consistent with the hot-plate assay, Y1^{-/-} mice showed a markedly reduced latency at all temperatures between 46 and 54 °C (Fig. 2b). Y1^{-/-} mice also displayed a

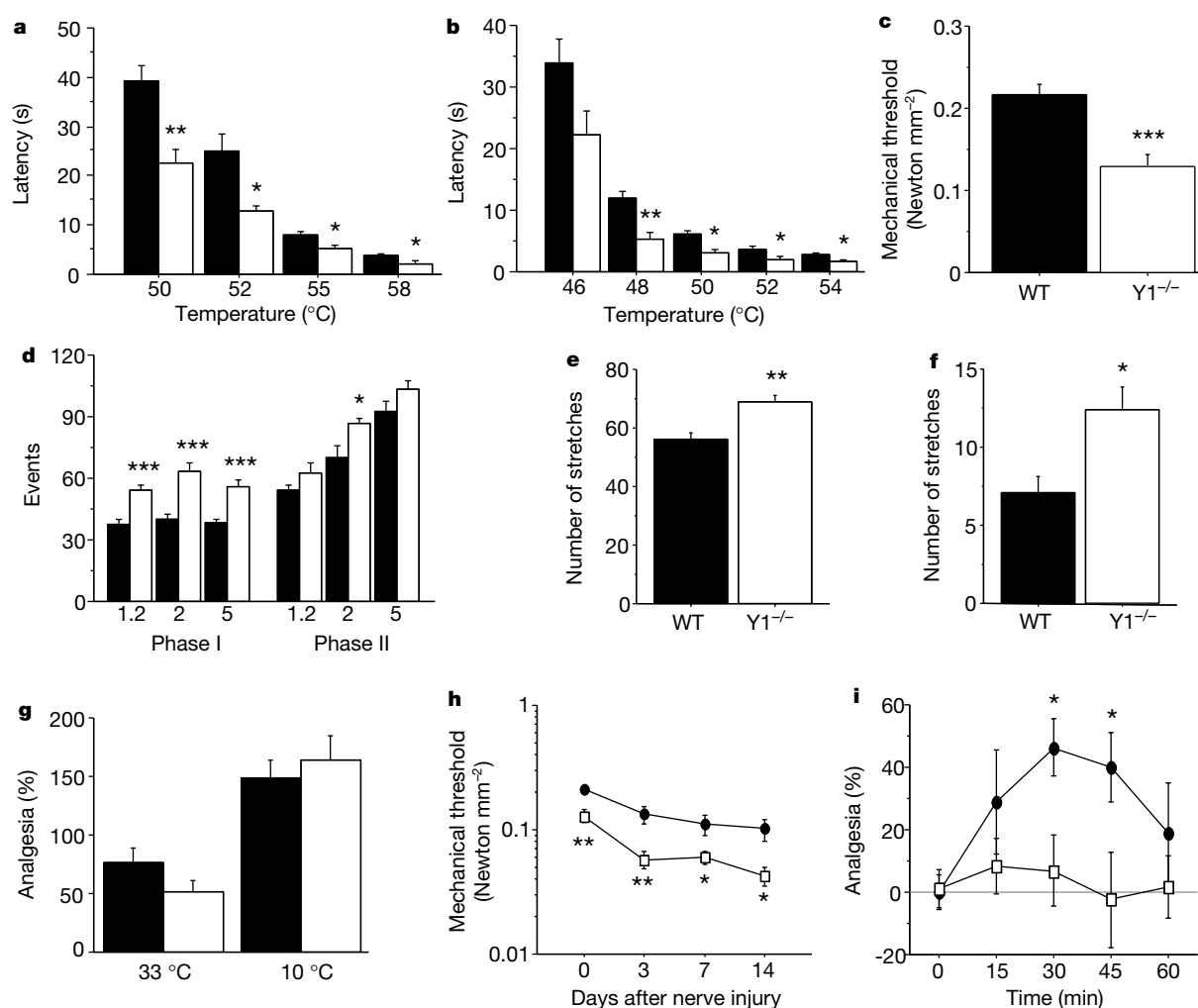


Figure 2 Cutaneous and visceral nociception of wild-type (black bars) and Y1^{-/-} (white bars) mice in the hot-plate, tail-flick, formalin, acetic acid, MgSO₄, von Frey hair and neuropathic pain assays, as well as in stress- and NPY-produced analgesia. **a**, Latency to shaking of hindpaw or jumping in the hot-plate assay. **b**, Tail-flick latency. **c**, Mechanical threshold assayed by von Frey hairs. **d**, Measurement of the number of events (lifting, shaking, licking and biting of the injected paw) in the formalin assay. The numbers on the x axis indicate the concentration in per cent of formalin administered subcutaneously. **e**, **f**,

Visceral pain response (abdominal stretching) produced by intraperitoneal injection of diluted acetic acid (**e**), or MgSO₄ (**f**). **g**, Stress-induced analgesia in the hot-plate assay. **h**, Development of mechanical allodynia of wild-type and Y1^{-/-} mice in a chronic pain model. **i**, Analgesic response to tail-flick following an intrathecal injection of NPY. Data are presented as percentage analgesia. All data are mean $\pm$ s.e.m.; analysis was done by unpaired Student's *t*-test (**a–g**, **i**) or two-tailed Mann–Whitney *U*-test (**h**). Asterisk, $P < 0.05$; two asterisks $P < 0.01$; three asterisks, $P < 0.001$.

significant decrease in mechanical threshold, indicating mechanical hypersensitivity (Fig. 2c).

We then examined a role for the NPY Y1 receptor in chemical nociception. During the first phase of the formalin assay, which provides a measure of the acute pain mediated by direct chemical activation of C-fibres, the nociceptive behaviour was augmented by 44, 60 and 46% at 1.2, 2 and 5% formalin injected, respectively, in $Y1^{-/-}$ mice (Fig. 2d). The second phase of nociception was not consistently altered by an absence of the NPY Y1 receptor. In two models of visceral pain, one that is secondary to an inflammatory response (acetic acid) and one that induces immediate pain independent of inflammation ($MgSO_4$)¹¹, we also found significantly increased pain behaviour in $Y1^{-/-}$ mice (Fig. 2e, f). Combined, the above results indicate that NPY might be acting on Y1 containing polymodal nociceptive afferents (C-fibres). Consistent with our results, these neurons span different modalities, and mediate pain transmission from both visceral and cutaneous tissues.

Stress induces analgesia through endogenous opioid- and non-opioid-dependent mechanisms¹². We tested whether these pathways interact with NPY-dependent analgesia by letting the mice swim in water at 10 °C, which produces a non-opioid, NMDA-dependent analgesia, and at 33 °C, which results in opioid-dependent analgesia^{13,14}. After a swim at 10 or 33 °C, the development of

analgesia assayed by hot plate was similar in wild-type and $Y1^{-/-}$ mice (Fig. 2g). These data suggest that the Y1 receptor is not a critical component in stress-induced analgesia.

The role of NPY in neuropathic pain is incompletely defined and a number of conflicting results with regards to possible NPY receptors involved have been reported^{4,15}. We tested the physiological role of the Y1 receptor in a model of neuropathic pain. A partial sciatic nerve ligation resulted in mechanical allodynia (sensitization to mechanical stimuli) in wild-type mice (37% and 52% increase in sensitivity at 3 and 14 days after nerve injury, respectively, as compared with day 0; Fig. 2h). As indicated before, the basal threshold of mechanical sensitivity was significantly decreased in non-lesioned $Y1^{-/-}$ mice. Despite this, the mechanical allodynia caused by nerve damage was significantly increased in $Y1^{-/-}$ mice as compared with wild-type mice (55% and 67% increase in sensitivity at 3 and 14 days after nerve injury, respectively, as compared with day 0; $P < 0.01$ for the slopes of the curves between day 0 and 14 after nerve injury).

Pharmacological NPY-induced analgesia to thermal stimuli following spinal delivery is well documented^{2,16}. To identify the receptor involved in the pharmacological effects of intrathecally administered NPY, we injected NPY (10 μ g) in the spinal cord of $Y1^{-/-}$ mice and measured heat sensitivity. The antinociceptive effect

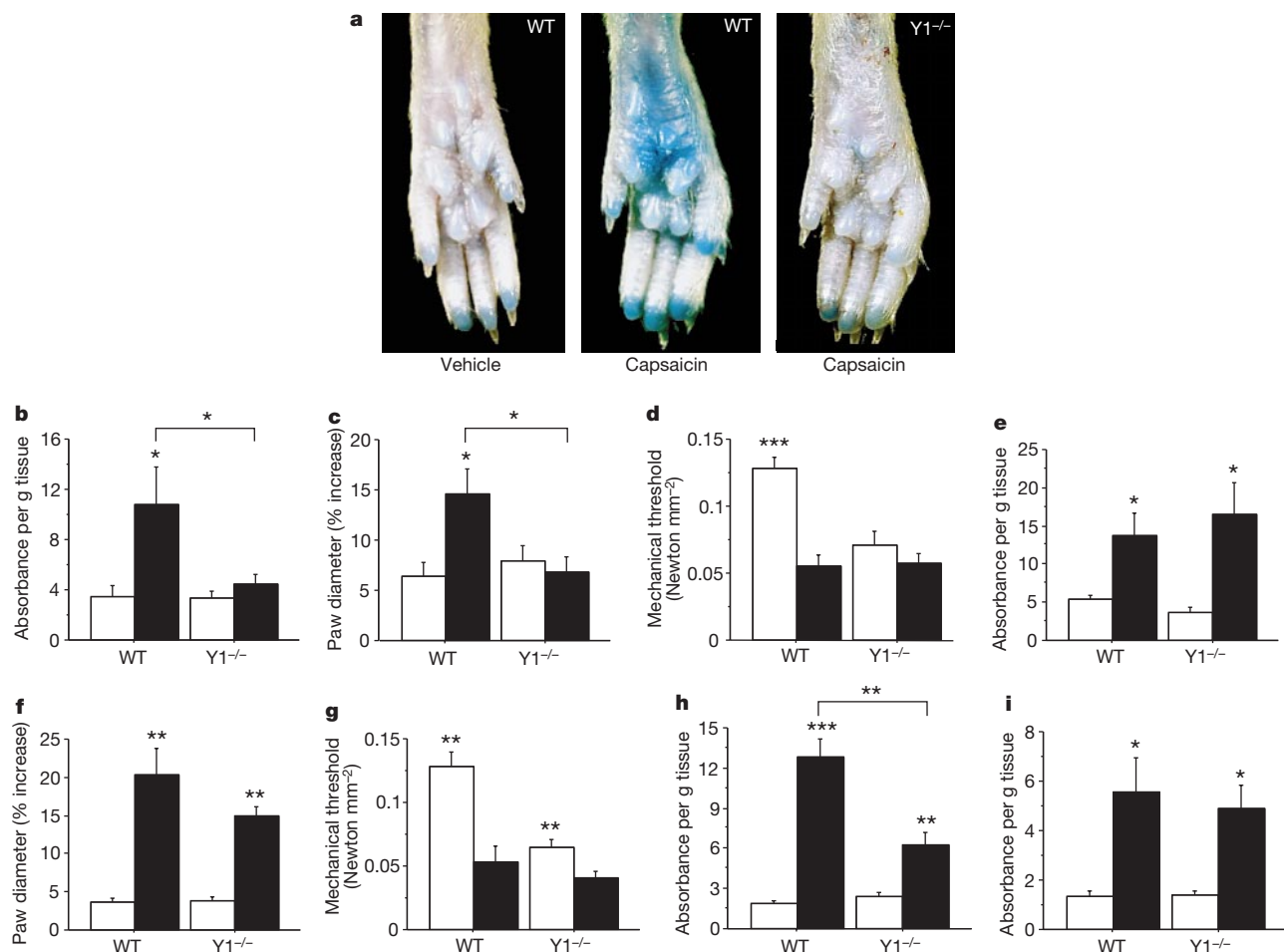


Figure 3 Neurogenic and non-neurogenic inflammation in wild-type and $Y1^{-/-}$ mice. **a**, Paws of wild-type and $Y1^{-/-}$ mice 30 min after injection of capsaicin (neurogenic inflammation) or vehicle. **b**, Quantification of Evans blue extravasation after capsaicin or vehicle injection. **c**, Percentage of paw diameter increase of vehicle and capsaicin injected paws. **d**, Mechanical sensitization before and after capsaicin-induced inflammation. **e**, **f**, Evans blue extravasation (**e**) and paw diameter (**f**) 4 h after carrageenan (non-neurogenic) induced inflammation in the wild-type and $Y1^{-/-}$ mice as indicated.

g, Mechanical sensitization 3 h after carrageenan induced inflammation. **h**, Quantification of Evans blue extravasation of the paws 30 min after mustard oil administration. **i**, Quantification of Evans blue extravasation 30 min after vehicle or SP administration. In all experiments open bars are the vehicle control side and black bars the experimental side. All data are mean $\pm$ s.e.m. Statistical analysis was performed by unpaired Student's *t*-test. Asterisk, $P < 0.05$; two asterisks, $P < 0.01$; three asterisks, $P < 0.001$.

of NPY on the spinal cord was completely abolished in $Y1^{-/-}$ mice (Fig. 2i). Thus, the $Y1$ receptor is exclusively responsible for the analgesic effects of centrally delivered NPY.

Inflammation is caused by both a neurogenic and a non-neurogenic component¹⁷. Neurogenic inflammation does not occur in the denervated human skin, can be prevented by a nerve block in rats^{17,18}, and is mediated by a peripheral release of SP/neurokinin A¹¹. We tested whether the $Y1$ receptor could participate in inflammation. A subcutaneous injection of capsaicin, which induces neurogenic inflammation, led to a marked inflammation seen by an increased paw diameter, plasma extravasation and hyperalgesia in wild-type mice. Unexpectedly, $Y1^{-/-}$ mice displayed no overt or quantitative sign of plasma extravasation, increase in paw diameter and hyperalgesia (Fig. 3a–d). Neither rectal and paw skin temperature nor heart rate differed between any of the groups before and after capsaicin administration (data not shown). Both wild-type and $Y1^{-/-}$ mice exhibited an increased blood flow in the paw as a response to capsaicin ($181.4 \pm 31.3\%$ and $245.5 \pm 71.4\%$, respectively) indicating that the $Y1$ receptor is not influencing capsaicin-induced vasodilation. In contrast to capsaicin, mustard oil induces inflammation that is largely, but not exclusively, dependent on a neurogenic component including the release and pro-inflammatory effects of SP¹⁹. $Y1^{-/-}$ mice showed markedly reduced plasma extravasation in response to mustard oil as compared with wild-type mice. However, as the increase in plasma extravasation was significant (although at a lower level), some components of the effects of mustard oil, probably those of non-neurogenic origin, were intact in $Y1^{-/-}$ mice (Fig. 3h). In contrast to neurogenic inflammation, there was a similar increase of plasma extravasation, paw diameter and sensitisation after non-neurogenic inflammation induced by carrageenan²⁰ in $Y1^{-/-}$ mice as in wild-type mice (Fig. 3e–g).

Because capsaicin and, to a large extent, mustard-oil-induced inflammation depend on the integrity of primary C-fibre afferent release of SP⁸, we determined whether $Y1$ is required before or after SP release in the sequence of events leading to inflammation by injecting SP in the paw. SP caused a similar inflammatory response in $Y1^{-/-}$ mice as it did in wild-type mice (Fig. 3i). SP receptor immunoreactivity was intact in the skin of wild-type and $Y1^{-/-}$ mice, and was found occasionally in nerve endings in dermis and in scattered cells throughout dermis corresponding to mast cells, similar to wild-type mice (data not shown). Our results are therefore consistent with $Y1$ being required for capsaicin-induced SP release.

Using EIA, we measured the quantity of total and released SP in the skin after capsaicin injection in wild-type and $Y1^{-/-}$ mice. Capsaicin caused a marked increase of released SP in the skin of wild-type mice, whereas it had no effect on SP release in $Y1^{-/-}$ mice (Fig. 4a). The lack of an increase of released SP was not caused by an overall reduction of SP peripherally because the quantity of total SP was similar in $Y1^{-/-}$ and wild-type mice ($2,269 \pm 524$ pg g⁻¹ tissue and $2,520 \pm 860$ pg g⁻¹ tissue, respectively). Furthermore, the number of SP immunoreactive nerve fibres in the dermis and epidermis was similar in wild-type and $Y1^{-/-}$ mice (5.7 ± 0.5 and 6.2 ± 0.5 per cm, respectively). In accordance with previous results, capsaicin led to a marked reduction of immunoreactive terminals in the epidermis of wild-type mice (from 3.0 ± 0.3 to 1.8 ± 0.2 per cm; $P < 0.01$, Student's *t*-test), possibly by a loss of immunoreactivity that was due to increased release. In contrast, $Y1^{-/-}$ mice displayed no reduction in SP-immunoreactive terminals (3.1 ± 0.3 and 2.5 ± 0.6 per cm, respectively). These results indicate that the absence of neurogenic inflammation in $Y1^{-/-}$ mice is caused by the requirement of $Y1$ activation for SP release. The persistent vasodilation in these mice might be caused by a normal release of a calcitonin-gene-related peptide, which induces vasodilation but not extravasation. Furthermore, because close to one order of magnitude less SP is required for vasodilation than for plasma leakage²¹, a small residual SP release in these mice might also cause this phenotype.

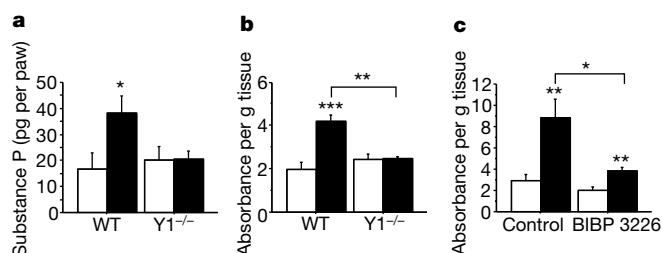


Figure 4 Measurement of SP release by capsaicin administration in the skin by EIA and effects of $Y1$ agonist and antagonist in inflammation-induced plasma extravasation. **a**, Released SP in vehicle and capsaicin injected skin. **b**, Evans blue extravasation 30 min after NPY $Y1$ receptor agonist [Leu^{31} -Pro³⁴]-NPY or vehicle injection intraplantarly. **c**, Capsaicin-induced Evans blue extravasation in wild-type mice in the presence or absence of NPY $Y1$ receptor selective antagonist BIBP 3226. In all experiments open bars are the vehicle control side and black bars the experimental side. All data are mean $\pm$ s.e.m. and statistical analysis was performed by unpaired Student's *t*-test. Asterisk, $P < 0.05$; two asterisks, $P < 0.01$; three asterisks, $P < 0.001$.

We challenged our results in wild-type mice by injecting the $Y1$ agonist [Leu^{31} -Pro³⁴]-NPY and by using a $Y1$ receptor antagonist. [Leu^{31} -Pro³⁴]-NPY efficiently caused plasma extravasation in wild-type mice, and consistent with its specificity for the $Y1$ receptor, no response was seen in $Y1^{-/-}$ mice (Fig. 4b). Administration of the $Y1$ antagonist BIBP 3226 before intraplantar injection of capsaicin markedly reduced plasma extravasation in wild-type mice (Fig. 4c), showing that the results on the $Y1^{-/-}$ mice are directly caused by the absence of $Y1$ signalling. We therefore conclude that the NPY $Y1$ receptor is both required and sufficient to induce neurogenic inflammation by controlling SP release, and that a $Y1$ antagonist could provide an effective strategy for the treatment of neurogenic inflammatory diseases.

We have shown that the $Y1$ receptor has an essential antinociceptive role during pain transmission in many modalities, including thermal, chemical and mechanical, from both cutaneous and visceral tissues, as well as in neuropathic pain. NPY or another $Y1$ receptor ligand²² may mediate antinociception by reducing SP and excitatory neurotransmitter release from primary C-fibre afferents^{3,23,24} and/or by inhibiting postsynaptically the SP-receptor-expressing projection neurons of the spinal cord^{25,26}. Consistent with the observation that NPY does not modulate pain transmission only through a presynaptic regulation of SP release, the nociceptive phenotype of the $Y1^{-/-}$ mice does not fully correlate with SP and SP receptor null mutant mice. SP receptor null mutant mice show, for example, a reduced stress-induced analgesia²⁷. We have also shown that $Y1$ receptor activation is both sufficient and required for neurogenic inflammation. Because mustard-oil-induced inflammation occurs independently of the vanilloid receptor that is activated by capsaicin¹⁹, our results indicate that activation of the $Y1$ receptor may be a shared and obligatory component in most, or all, neurogenic inflammatory conditions. □

Methods

$Y1$ gene targeting

Exon 2 of the $Y1$ gene was partially deleted and replaced by a IRES-*tau-lacZ* cassette also containing a neomycin-resistance gene driven by the PGK promoter and polyA (ETLN). A 0.7-kilobase DNA fragment 3' from the $Y1$ targeting construct was used as external probe. Homologous recombinant embryonic stem cell clones were injected to generate $Y1$ mutant 129SVXBalb/c hybrid mice. Mice were analysed by Southern blot and PCR using the primers 5'-ATCAAAATCTGACCGACGAG-3', 5'-CATGATGTTGATTCGCTTGG-3' and 5'-GCAGCCTCTGTTCCACATACA-3'. Standard procedures were used for northern blot analysis, and 60 μ g per sample of total RNA from adult brain was analysed.

Physiological studies

Heart rate and body temperature was measured as previously described²⁸. Skin blood flow and temperature was monitored with a thermal probe (Physitemp bat-12) subcutaneously

inserted in the plantar region of the paw and a Laser doppler flowmeter probe applied to the plantar surface (Permedid PF2B) was used to measure basal and capsaicin-evoked changes. Student's *t*-test was used and differences were considered significant at $P < 0.05$.

Behavioural studies

Between 6 and 13 adult male F2-3 129SV×Balb/c mice of each genotype were used for all studies. Thermal¹¹, mechanical¹¹, chemical¹¹, visceral¹¹ sensitivity, stress-induced analgesia²⁷ and neuropathic pain^{29,30} was assessed as described previously. Tail-flick latency (by directing a concentrated light beam to the tail of the mouse) was monitored before and after intrathecal injection of 10 µg NPY. For Evans blue plasma extravasation the following were used: capsaicin, 3 µg in 10 µl (Sigma; dissolved in 5% ethanol, 5% Tween-80 and 90% saline); 1% carrageenan (Sigma; dissolved in saline); SP, 50 pmol per paw (Sigma; dissolved in saline); 5% mustard oil (Fluka; dissolved in mineral oil); NPY Y1 receptor agonist [Leu³¹-Pro³⁴]-NPY, 10 µg per paw (Calbiochem; dissolved in saline and 5% acetic acid); and NPY Y1 receptor selective antagonist BIBP 3226, 10 mg kg⁻¹ in 10 ml kg⁻¹ (American Peptide; dissolved in saline and 5% acetic acid). Briefly, mice were anaesthetized and injected intravenously with Evans blue (50 mg kg⁻¹) into the jugular vein. The above agents were injected into one paw of the animal except for the Y1 receptor selective antagonist BIBP 3226, which was injected intravenously 10 min before injection into the paw. The other paw was injected with vehicle. After 30 min the plantar skin of the paw was removed, dried of excess liquid, weighed and incubated in formamide for 24 h at 56 °C. Extravasated Evans blue was measured by spectrophotometer at 620 nm. Mechanical sensitivity was determined before, 30 min and 3 h after capsaicin and carrageenan administration, respectively. Carrageenan inflammation was induced similarly but extravasation was measured after 4 h. The paw diameter was measured before and after capsaicin, carrageenan or vehicle administration using a spring-loaded calliper.

Immunohistochemistry

Wild-type and Y1^{-/-} mice were perfused with 4% paraformaldehyde (for SP receptor immunohistochemistry, mice were perfused with 4% paraformaldehyde and 12.5% picric acid 10 min after capsaicin injection into hindpaw), and the spinal cord and dorsal root ganglia were sectioned coronally (15 µm in thickness). Capsaicin was injected intradermally into dorsal skin of mice. After 10 min the skin was removed, postfixed and sectioned as above. Immunohistochemistry was done as described²⁸ using anti-β-galactosidase (1:200 dilution, ICN/Cappel), rabbit anti-SP (1:5,000 dilution, Chemicon), guinea-pig anti-SP (1:200 dilution, Peninsula Lab.), rhodamine-conjugated bandersea simplicifolia lectin I (Isolectin B4; 1:100 dilution, Vector), anti-NPY (1:200 dilution, Peninsula Lab.), and rhodamine or FITC-conjugated secondary antisera (Jackson). For SP receptor immunohistochemistry, sections were incubated for 30 min in PBS, 50% methanol, and 0.6% H₂O₂ before incubation in 10% goat serum. The antiserum (Chemicon 1:2,000) was used in the fluorescein TSA fluorescence system (NEN). β-galactosidase histochemical staining was done as described²⁸.

EIA

Capsaicin or saline was injected into the paw of wild-type or Y1^{-/-} mice. After 10 min, the paw was removed and the skin was cut open and washed in PBS and 0.1% BSA for 10 min. The skin was then dried, weighed, transferred to a new container, and frozen. The liquid was centrifuged at 4,000 r.p.m. for 15 min. Supernatant was transferred to a new tube, weighed and frozen. The lumbar part of spinal cord was removed, weighed and frozen. The samples were then assayed for SP according to the manufacturer's instructions using SP high sensitivity EIA kit (Peninsula Lab.).

Received 29 September; accepted 13 November 2000.

- Munglani, R., Hudspeth, M. J. & Hunt, S. P. The therapeutic potential of neuropeptide Y. Analgesic, anxiolytic and antihypertensive. *Drugs* **52**, 371–389 (1996).
- Hua, X. Y. *et al.* The antinociceptive effects of spinally administered neuropeptide Y in the rat: systematic studies on structure-activity relationship. *J. Pharmacol. Exp. Ther.* **258**, 243–248 (1991).
- Duggan, A. W., Hope, P. J. & Lang, C. W. Microinjection of neuropeptide Y into the superficial dorsal horn reduces stimulus-evoked release of immunoreactive substance P in the anaesthetized cat. *Neuroscience* **44**, 733–740 (1991).
- Hudspeth, M. J. & Munglani, R. Neuropeptide Y: friend or foe. *Eur. J. Pain* **3**, 3–6 (1999).
- Pedrazzini, T. *et al.* Cardiovascular response, feeding behavior and locomotor activity in mice lacking the NPY Y1 receptor. *Nature Medicine* **4**, 722–726 (1998).
- Naveilhan, P., Neveu, I., Arenas, E. & Ernfor, P. Complementary and overlapping expression of Y1, Y2 and Y5 receptors in the developing and adult mouse nervous system. *Neuroscience* **87**, 289–302 (1998).
- Zhang, X. *et al.* Localization of neuropeptide Y Y1 receptors in the rat nervous system with special reference to somatic receptors on small dorsal root ganglion neurons. *Proc. Natl Acad. Sci. USA* **91**, 11738–11742 (1994).
- Snider, W. D. & McMahon, S. B. Tackling pain at the source: new ideas about nociceptors. *Neuron* **20**, 629–632 (1998).
- Allen, B. J. *et al.* Primary afferent fibers that contribute to increased substance P receptor internalization in the spinal cord after injury. *J. Neurophysiol.* **81**, 1379–1390 (1999).
- Mantyh, P. W. *et al.* Receptor endocytosis and dendrite reshaping in spinal neurons after somatosensory stimulation. *Science* **268**, 1629–1632 (1995).
- Cao, Y. Q. *et al.* Primary afferent tachykinins are required to experience moderate to intense pain. *Nature* **392**, 390–394 (1998).
- Watkins, L. R. & Mayer, D. J. Multiple endogenous opiate and non-opiate analgesia systems: evidence of their existence and clinical implications. *Ann. N. Y. Acad. Sci.* **467**, 273–299 (1986).
- Rubinstein, M. *et al.* Absence of opioid stress-induced analgesia in mice lacking β-endorphin by site-directed mutagenesis. *Proc. Natl Acad. Sci. USA* **93**, 3995–4000 (1996).

- Marek, P., Mogil, J. S., Sternberg, W. F., Panocka, I. & Liebeskind, J. C. N-methyl-D-aspartic acid (NMDA) receptor antagonist MK-801 blocks non-opioid stress-induced analgesia. II. Comparison across three swim-stress paradigms in selectively bred mice. *Brain Res.* **578**, 197–203 (1992).
- Zhang, X., Wiesenfeld-Hallin, Z. & Hokfelt, T. Effect of peripheral axotomy on expression of neuropeptide Y receptor mRNA in rat lumbar dorsal root ganglia. *Eur. J. Neurosci.* **6**, 43–57 (1994).
- Xu, X. J., Hao, J. X., Hokfelt, T. & Wiesenfeld-Hallin, Z. The effects of intrathecal neuropeptide Y on the spinal nociceptive flexor reflex in rats with intact sciatic nerves and after peripheral axotomy. *Neuroscience* **63**, 817–826 (1994).
- Jancso, N., Jancso-Gabor, A. & Szolcsanyi, J. Direct evidence for neurogenic inflammation and its prevention by denervation and by pretreatment with capsaicin. *Br. J. Pharmacol.* **31**, 138–151 (1967).
- Jancso, N., Jancso-Gabor, A. & Szolcsanyi, J. The role of sensory nerve endings in neurogenic inflammation induced in human skin and in the eye and paw of the rat. *Br. J. Pharmacol.* **33**, 32–41 (1968).
- Inoue, H., Asaka, T., Nagata, N. & Koshihara, Y. Mechanism of mustard oil-induced skin inflammation in mice. *Eur. J. Pharmacol.* **333**, 231–240 (1997).
- Semos, M. L. & Headley, P. M. The role of nitric oxide in spinal nociceptive reflexes in rats with neurogenic and non-neurogenic peripheral inflammation. *Neuropharmacology* **33**, 1487–1497 (1994).
- Lembeck, F. & Holzer, P. Substance P as neurogenic mediator of antidromic vasodilation and neurogenic plasma extravasation. *Naunyn-Schmiedelbergs Arch. Pharmacol.* **310**, 175–183 (1979).
- Shi, T. J., Zhang, X., Berg, O. G., Erickson, J. C., Palmiter, R. D. & Hokfelt, T. Effect of peripheral axotomy on dorsal root ganglion neuron phenotype and autotomy behaviour in neuropeptide Y-deficient mice. *Regul. Pept.* **25**, 161–173 (1998).
- Wilcox, G. L. *Excitatory Neurotransmitters and Pain* (eds Bond, M. R., Charlton, J. R. & Woolf, C. J.) 97–117 (Elsevier, New York, 1991).
- Walker, M. W., Ewald, D. A., Perney, T. M. & Miller, R. J. Neuropeptide Y modulates neurotransmitter release and Ca²⁺ currents in rat sensory neurons. *J. Neurosci.* **8**, 2438–2446 (1988).
- Polgar, E., Shehab, S. A., Watt, C. & Todd, A. J. GABAergic neurons that contain neuropeptide Y selectively target cells with the neurokinin 1 receptor in laminae III and IV of the rat spinal cord. *J. Neurosci.* **19**, 2637–2646 (1999).
- Parker, D. *et al.* Co-localized neuropeptide Y and GABA have complementary presynaptic effects on sensory synaptic transmission. *Eur. J. Neurosci.* **10**, 2856–2870 (1998).
- De Felipe, C. *et al.* Altered nociception, analgesia and aggression in mice lacking the receptor for substance P. *Nature* **392**, 394–397 (1998).
- Naveilhan, P. *et al.* Normal feeding behavior, body weight and leptin response require the neuropeptide Y Y2 receptor. *Nature Medicine* **5**, 1188–1193 (1999).
- Seltzer, Z., Dubner, R. & Shir, Y. A novel behavioral model of neuropathic pain disorders produced in rats by partial sciatic nerve injury. *Pain* **43**, 205–218 (1990).
- Malmberg, A. B. & Basbaum, A. I. Partial sciatic nerve injury in the mouse as a model of neuropathic pain: behavioral and neuroanatomical correlates. *Pain* **76**, 215–222 (1998).

Acknowledgements

We thank L. Klevenvall-Fridvall for technical assistance and L. Johansson for secretarial assistance. This research was supported by the Swedish Medical Research Council the Biotechnology Program of the European Union and the Swedish Cancer Society.

Correspondence and requests for material should be addressed to P.E. (e-mail: patrik@cajal.mbb.ki.se).

Role for sperm in spatial patterning of the early mouse embryo

Karolina Piotrowska & Magdalena Zernicka-Goetz

Wellcome/CRC Institute and Department of Genetics, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK

Despite an apparent lack of determinants that specify cell fate, spatial patterning of the mouse embryo is evident early in development. The axis of the post-implantation egg cylinder can be traced back to organization of the pre-implantation blastocyst¹. This in turn reflects the organization of the cleavage-stage embryo and the animal-vegetal axis of the zygote^{2,3}. These findings suggest that the cleavage pattern of normal development may be involved in specifying the future embryonic axis; however, how and when this pattern becomes established is unclear. In many animal eggs, the sperm entry position provides a cue for embryonic patterning^{4–6}, but until now no such role has been found in mammals. Here we show that the sperm entry position predicts the plane of initial cleavage of the mouse egg and can define embryonic and abembryonic halves of the future blasto-

cyst. In addition, the cell inheriting the sperm entry position acquires a division advantage and tends to cleave ahead of its sister. As cell identity reflects the timing of the early cleavages, these events together shape the blastocyst whose organization will become translated into axial patterning after implantation. We present a model for axial development that accommodates these findings with the regulative nature of mouse embryos.

Early mouse development is highly regulative and consequently axis formation has been treated as being independent of any aspect of organization of the egg. Thus, there has been no incentive to study correlations between the sperm entry position (SEP) and developmental patterns. As this viewpoint has now been questioned¹⁻³, we wanted to determine whether there are any developmental consequences of sperm entry on the subsequent patterning of the mouse embryo.

We therefore developed an enduring marker for the SEP. A transitory protrusion, the fertilization cone, appears at the SEP above the condensed chromatin of the sperm, but is not detectable after the sperm head has decondensed and formed the pronucleus (Fig. 1a, b). We found that the SEP could be marked by phytohaemagglutinin-coated fluorescent beads introduced under the zona pellucida and placed in contact with the plasma membrane at the fertilization cone. Beads remained adhered to the plasma membrane throughout subsequent development without undergoing endocytosis.

To determine whether any fluidity of the plasma membrane might preclude the use of such beads as positional markers, we first ascertained whether they would maintain their relative positions with respect to the second polar body (PB), a marker of the animal pole of the egg until the blastocyst stage^{2,3} (Fig. 1). We found that when beads were positioned near to the PB they remained in this position in 82% (47/57) of embryos throughout the two-cell, morula and blastocyst stages. Similarly, when placed opposite the PB, beads retained their position relative to the PB throughout these stages in 85% (41/48) of zygotes. In the two groups, a small proportion of zygotes (18% and 15% respectively) showed relative movement of the beads away from their original position. This might be explained by the previously observed rotation of blastomeres after cleavage divisions⁷ and was seen as a slight displacement of the bead at the two-cell stage. After such displacement, however, the new position of the bead was subsequently maintained with respect to the PB until the blastocyst stage. Therefore, when the surface of the egg is marked in this way, its features (the PB and the bead) retain their relative positions throughout pre-implantation development in most embryos, and thus can be used as landmarks for the animal pole of the zygote and the site of sperm entry, respectively.

We used beads to mark the SEPs of newly fertilized zygotes to follow their positions in the subsequent development of the embryo. Differential development is first seen in the two-cell

embryo in the form of asynchrony of the second cleavage divisions. To determine whether there was any relationship between the order of cleavage and the SEP, we examined the position of the bead in three-cell embryos in which one of the second cleavages had just occurred. We found that in 75% of a group of 92 three-cell embryos the bead marked one of the two smaller cells, indicating that the blastomere inheriting the SEP had divided ahead of its sister (Fig. 1c). This suggests that the incoming sperm can provide information to its host egg cytoplasm that influences the order in which cells divide.

We then allowed these embryos to develop into blastocysts to determine whether the SEP would come to lie in any specific region, scoring its position first with respect to the animal-vegetal (A-V) axis. The sperm can enter the egg at any point on its surface with the exception of an area above the metaphase II spindle. This is reflected in the distribution of the SEP marker along the A-V axis in a group of 73 zygotes in which sperm entered in the animal third in 16% of cases, the equatorial third in 37%, and the vegetal third in 47%. From 50 blastocysts that developed from these zygotes with an intact PB on the A-V axis, 16% had the bead in the animal third, 32% in the equatorial region, and 52% in the vegetal third. This corresponds to the proportion of zygotes that had their SEPs positioned in these regions, confirming that landmarks on the

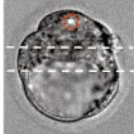
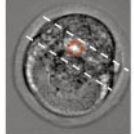
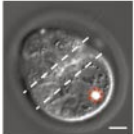
	Position of the bead on Em-Ab axis of blastocyst		
Initial position of the bead in the zygote			
a Sperm entry position 	10/57 (17%)	42/57 (74%)	5/57 (9%)
b Vegetal pole (control) 	22/48 (46%)	22/48 (46%)	4/48 (8%)
c Equatorial region (control) 	11/23 (48%)	8/23 (35%)	4/23 (17%)

Figure 2 The SEP predicts the Em-Ab axis of the blastocyst. **a**, Zygotes were marked with beads at the SEP and examined at the blastocyst stage. The position of the bead was scored depending on whether it lay in the embryonic (left), medial (middle), or abembryonic region (right). Representative micrographs of such blastocysts are shown in which the limits of the medial regions are indicated by dashed white lines. **b**, **c**, Controls in which the beads were placed either at the vegetal pole (**b**) or in the equatorial region (**c**), and scored as above.

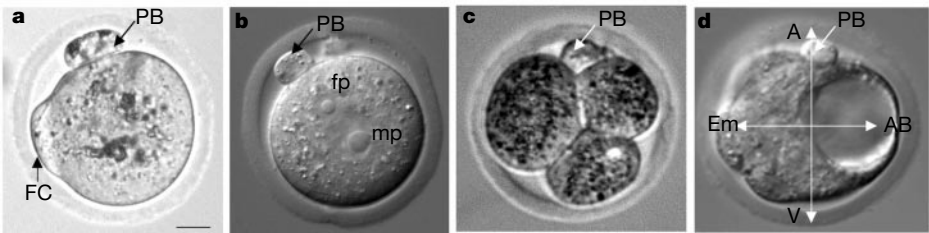


Figure 1 Landmarks in the development of the pre-implantation embryo. **a**, Egg shortly after fertilization, showing the fertilization cone and the PB. **b**, Egg after migration of the male pronucleus (mp) and female pronucleus (fp) towards the centre of the zygote. **c**, Three-cell-stage embryo in which one of the second cleavage divisions has taken place. One of the two small cells is marked with a bead that was placed on the zygote at

the SEP. **d**, Blastocyst shown in a comparable orientation to the eggs in **a-c** to indicate its two principal axes, A-V and Em-Ab. The A-V axis of the blastocyst lies along the medial region of the blastocyst tangential to the blastocyst cavity (blastocoel) and reflects the A-V axis of the zygote^{2,3}. The blastocoel itself occupies the abembryonic hemisphere. Scale bar in all figures, 20 μ m.

zygote retain their relative positions in the blastocyst.

We next analysed the distribution of the SEP along the embryonic–abembryonic (Em–Ab) axis of the blastocyst. This lies orthogonal to the A–V axis, and is morphologically asymmetric in that the blastocyst cavity occupies the abembryonic hemisphere (Fig. 1d). Analysis of the position of the beads showed that their distribution along this axis was remarkably non-random. To map the position of the bead on the blastocyst, we defined the medial region of the Em–Ab axis to be within three bead diameters (roughly 12.5% of the embryo diameter) below or above the blastocoel roof (Fig. 2) and found that in 74% of cases the bead came to lie in this region. This shows that in most embryos the SEP comes to lie near the border of the embryonic and abembryonic hemispheres.

To investigate the possibility that the bead tends to become located in this region as a result of some physical constraint bearing no relationship to the SEP, we analysed two sets of control experiments in which beads were placed on the surface of the zygote at positions that may or may not have included the SEP. The first set was from the experiment described above in which beads were placed opposite the PB of the zygote and followed to the blastocyst. We found that for the 48 embryos analysed, the bead lay in the embryonic region in 46% of cases, in the medial region in 46%, and in the abembryonic region in 8% (Fig. 2, control 2). We also used beads to mark the zygote surface at a random position within the equatorial region. In these experiments beads came to lie in the embryonic region in 48% of blastocysts, in the medial region in 35%, and in the abembryonic region in 17% (Fig. 2, control 3). Thus when beads are placed at sites without respect to the SEP, they show significantly different distributions (χ^2 test, $P < 0.001$, 2 degrees of freedom) with a reduced tendency to lie in the medial region of the blastocyst. Hence by marking the Em–Ab boundary, the SEP predicts the orientation of the Em–Ab axis with which it aligns orthogonally.

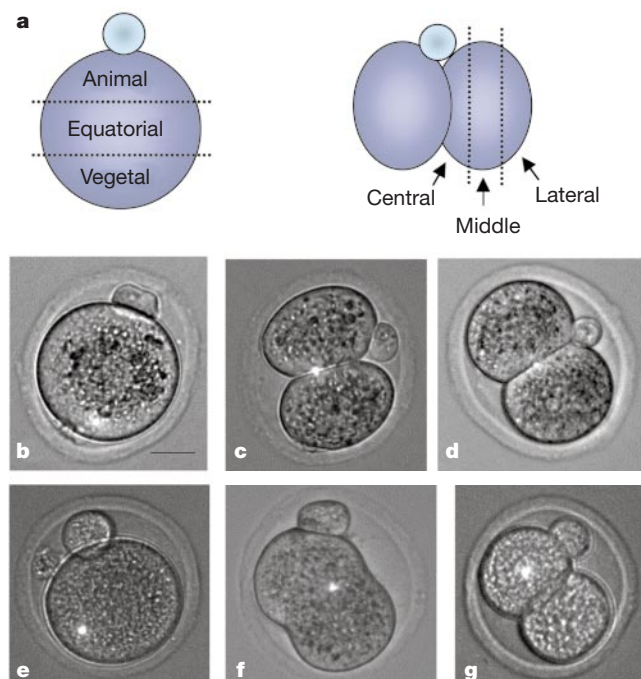


Figure 3 The SEP predicts the first cleavage plane. **a**, Sectors in which the position of the bead was scored in the zygote and the two-cell stage. **b–d**, Micrographs showing the first cleavage of an egg marked at its SEP. In this example, the bead is near the vegetal pole of the zygote (**b**), is at the furrow during cleavage (**c**), and remains on the furrow that separates the cells (**d**). **e–g**, A second example of an egg in which the bead is located at the position of the forming furrow during cleavage (**f**), but was displaced away from the furrow that later separates the two cells (**g**).

We considered whether this tendency of the SEP to lie in a particular region of the blastocyst might reflect some earlier developmental event such as the pattern of early cleavage. It is known that the first cleavage plane of the zygote is meridional, that is, parallel to the A–V axis of the zygote as defined by the PB at its animal pole. However, the orientation of this initial cleavage about the A–V axis has been thought to be random and not specified by any earlier developmental event³. To determine whether the SEP bears a relationship to this initial cleavage we followed the development of 178 zygotes. First we marked the SEP and estimated the position of the bead with respect to the A–V axis before the fertilization cone disappeared and confirmed its position when the male pronucleus had formed but not yet migrated to the egg centre. In the zygote we recorded the SEP in one of three sectors: animal, equatorial and vegetal (see Fig. 3).

Next we examined the position of the bead shortly after completion of the first division, and subsequently at a more advanced two-cell stage. At the two-cell stage, we further defined its position within one of three equal sectors: central, middle and lateral. As the first cleavage is meridional with respect to the PB, these sectors lie orthogonal to those in the zygote. If the SEP and the plane of the first cleavage were random with respect to each other, we would expect the bead to lie with equal probability (33%) in each of these three regions. However, we found that just after the first cleavage the SEP tended to be at or in close proximity to the furrow separating the cells. In 178 two-cell embryos analysed, the SEP was localized to the central zone in 60% of embryos, and in the middle and lateral regions in only 19% and 21%. When those embryos (107 of 178) that showed a central location of SEP were further analysed, roughly equal numbers had SEP originating from the animal (60%), equatorial (57%) and vegetal (67%) regions of the zygote. Beads

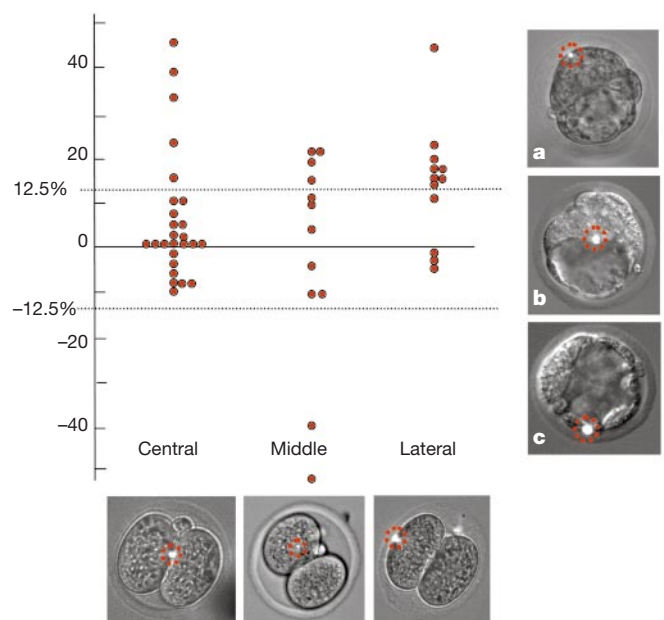


Figure 4 Relationship between the SEP at the two-cell stage and its fate at the blastocyst stage. Zygotes marked at their SEP were classified into three categories after their first cleavage, depending on whether the bead lay in the central, middle or lateral sectors of 1/2 blastomeres. Shown are representative micrographs of two-cell embryos in which the bead occupied either the boundary zone (**b**) or was localized to either the embryonic (**a**) or abembryonic pole (**c**). Zygotes were allowed to develop to the blastocyst and the final position of the bead was determined. The value plotted here is the distance of the bead from the border of the embryonic and abembryonic regions delineated by the roof of the blastocoel projected onto a flat plane and assessed as a percentage of the Em–Ab diameter of each blastocyst. The boundary zone is shown as the area within 12.5% on each side of the blastocoel roof.

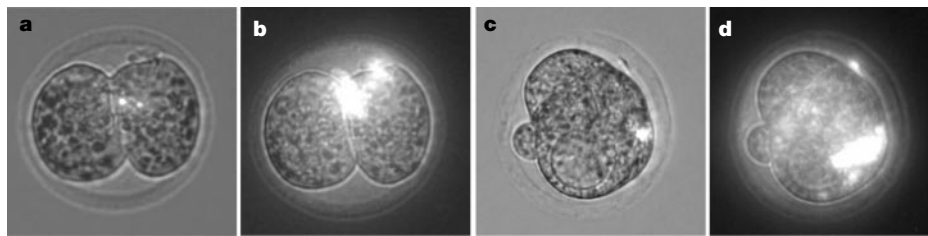


Figure 5 Two markers of the egg surface colocalize from the SEP in the zygote to the blastocyst. Zygotes were marked with beads (**a, c**) and ConA (**b, d**) at the SEP, and examined at the two-cell (**a, b**) and blastocyst (**c, d**) stages. Two markers colocalized at

the two-cell stage with the ConA extending up against the cleavage furrow. At the blastocyst stage, both markers are at the boundary zone between embryonic and abembryonic regions, with the beads lying in the area delimited by the labelled ConA.

that were placed randomly in the equatorial regions (not necessarily excluding the SEP) of 28 zygotes showed no significant tendency of localization to the cleavage furrow: 38% were in the central, 42% in the middle, and 20% in the lateral regions. These findings show that irrespective of the position of the SEP in the zygote, there is a strong tendency for the SEP to be localized to the central region near the cleavage plane of the two-cell embryo.

As some rotation of cells can occur after the first cleavage, we determined the extent to which the position of the furrow finally separating the two cells corresponded to the actual site of cytokinesis. By catching 37 embryos directly as the first cleavage division was in progress, we found the bead at or very close to the cleavage furrow in 78% (29/37) of cases (Fig. 3c, f). Unexpectedly, the bead appeared to indicate the plane of cleavage more accurately than did the PB, as only four of these embryos (11%) had their PBs near to the cleavage furrow. When these 37 embryos were observed 2–3 h later, there had been a slight reorientation of blastomeres such that the PB was now located in the furrow between the two cells. If the bead and the PB were attached to the same blastomere, this rotation was able to shift the bead slightly away from the furrow separating the cells (Fig. 3f, g). Direct observation of these embryos suggested that this rotation could occur within minutes of the completion of cytokinesis. Thereafter over the next several hours there was very little blastomere rotation. Thus, there is a strong correlation between the position of sperm entry and the first cleavage plane, and this is still evident even after some blastomeres have undergone rotation.

To determine the relationship between the SEP, the plane of the first cleavage and the Em–Ab axis, we cultured a set of embryos from the above experiments to the blastocyst stage. To minimize developmental disturbances we examined a subgroup of embryos that had been observed twice during their early development: initially to confirm the position of the bead, and then shortly after the first cleavage. A total of 50 such blastocysts were scored in three groups depending whether the SEP was located in the central (52%), middle (24%) or lateral (24%) sectors of the 1/2 blastomere. At the blastocyst stage, we defined the boundary between embryonic and abembryonic regions as a zone equivalent to 12.5% of the embryo diameter lying on either side of the blastocyst cavity roof (Fig. 4). In the main group where the SEP had been in the central sector at the two-cell stage, 81% of blastocysts had the SEP in this boundary zone. This compared to 50% of the group where the SEP had occupied the middle sector and 33% where it had been in the lateral sector at the two-cell stage. The number of embryos with the SEP positioned between 1/2 blastomeres after cleavage is completed appears to underestimate the proportion of zygotes where it is at or close to the first cleavage plane during actual division itself (as a result of blastomere rotation; see above). Thus, the proportion of embryos in the ‘central’ group could have been initially greater than that observed here.

As the bead used to mark the SEP labelled only a small area of the fertilization cone, we sought an independent method to mark a

larger patch of the egg surface around the SEP so that we could better relate this to the cleavage furrow. This was achieved by introducing a microdrop of rhodamine-labelled concanavalin A (ConA) solution onto the egg surface. We introduced both such a label and two fluorescein-labelled beads onto the fertilization cone and followed their development to the blastocyst stage in a series of 15 zygotes. In all cases we found that the two markers colocalized at the two-cell stage with the ConA extending up against the cleavage furrow. In 13 out of 15 blastocysts, both markers were at the Em–Ab border, and in all cases beads were in the area delimited by the labelled concanavalin (Fig. 5). Together, this shows that the SEP is a strong predictor of both the plane of first cleavage and the Em–Ab boundary of the blastocyst. We confirmed this by precisely marking the furrow separating blastomeres in two-cell stage embryos and found that in 78% (21/27) of cases the label came to lie against the Em–Ab border.

These studies show that the SEP defines a plane of cleavage that can separate the embryo into distinct halves, and is thus an important indicator of the future axial development of the embryo. This is an unexpected finding, not only because the first cleavage was previously thought to orientate randomly about the A–V axis, but also because it indicates that embryonic axes become specified at the earliest possible stage of zygotic development. The latter was not anticipated because early blastomeres are totipotent, allowing the embryo to regulate its development to overcome perturbations in its organization^{2,8–10}. However, totipotency of both two-cell stage blastomeres and their ability to contribute to inner cell mass (ICM) and trophoblast lineages¹¹ does not preclude them from following a specific developmental pathway. Division of the embryo into embryonic and abembryonic hemispheres by the first cleavage predicts a model in which descendants of one cell would normally contribute mainly to mural trophoblast and ICM that will become primitive endoderm, whereas the other blastomere would develop predominantly into the polar trophoblast and ICM destined to become epiblast (Fig. 6). This could be explained if the first cleavage divided the zygote into two halves having different constituents, a situation for which there is currently no proof.

But there is an alternative possibility in which regulating the pattern of the cleavage divisions may provide a way to specify embryonic axes that is not determinative but equips the embryo with regulative flexibility. It might not be the cleavage plane *per se* that separates cells with differing destinies, but rather that the fate of a particular blastomere is a direct consequence of having inherited the SEP. In this case the identity of each blastomere arising from the first cleavage would be a consequence of which one divides first. In perturbed development, we propose that regulation could be achieved if developmental processes followed the cue of whichever cell acquired such a division advantage through stochastic events.

A number of independent studies have suggested that earlier-dividing cells become preferentially incorporated into the ICM^{7,12–14}. Alternatively, early-dividing cells show a greater tendency

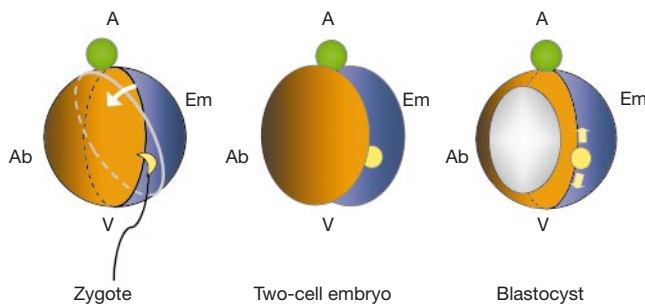


Figure 6 A model to show the developmental consequences of the first cleavage division defined by the SEP. The plane of first cleavage bisects the zygote (into copper and mauve hemispheres) passing close to the PB (green) and the SEP (yellow). The angle of this plane can vary slightly (direction of the white arrow) depending on whether it passes on one or other side of either the PB or the SEP (grey disc) or whether both PB and SEP together fall on the same side of the plane (not shown). This variability is transmitted to later stages. The corresponding position of the bead marking the SEP (yellow) is shown, together with the distinction of the abembryonic (Ab; copper) or embryonic (Em; mauve) lineages at the two-cell and blastocyst stages. The identity of cells could also be influenced as a result of the division advantage given to the blastomere inheriting the SEP. Although the SEP is shown to fall on the Em side of the cleavage plane in this drawing, there is currently no rigorous evidence to indicate a preference for either the Em or Ab side. The yellow arrows indicate that the fate of the SEP can vary in position along the Em–Ab boundary of the blastocyst, depending on the point at which the sperm penetrates the egg.

to be associated with the nascent blastocoel and thus the abembryonic pole¹⁵. Together, the directed orientation of the initial cleavage and predisposition of the earlier dividing cells to differentiate first may explain both their positioning on the Em–Ab boundary and why the ICM and blastocoel intersect at this site. Previous lineage tracing of the fates of one out of eight cell blastomeres located either adjacent to or opposite the PB should have revealed a tendency of such cells to lie on one or other side of the Em–Ab boundary². Yet a substantial proportion extended between these regions. This might support a model in which cells with a division advantage have a greater role in setting identity. But it is difficult to position the Em–Ab boundary by lineage tracing because, although clones do at first stay coherent, at the later preimplantation stages they can become disturbed in the blastocyst when polar trophectoderm becomes recruited to mural trophectoderm^{16–18}. Moreover, the Em–Ab border as defined by the first cleavage plane might not be readily seen from morphological features.

Another model predicts that the border would not be fixed by the position of the blastocoel roof, but would be located parallel to it within the ICM showing slight variation depending whether the first cleavage passed on one or other side of the SEP or PB (Fig. 6). Thus, the position of the SEP and PB may determine the orientation but not the polarity of the Em–Ab axis, the latter potentially being subject to the position of the SEP on either side of the first cleavage plane. Indeed, it is not clear whether the blastomere inheriting the SEP has a distinct embryonic or abembryonic fate. It will be a future objective to understand the cellular and molecular mechanism of the effect exerted by the SEP, its developmental timing, and whether it occurs at the surface or within the egg.

We conclude that two axes of the blastocyst become specified in the single-cell embryo. One of these is defined by the animal pole and the second, the Em–Ab axis, relates to the SEP. These axes are initially not fixed and can be re-established if development is perturbed^{2,19}. Directing blastocyst organization by the plane of cleavage and cellular identity by the order of cleavage may offer an interpretation of the regulative events that occur following perturbation of development. In normal development the orientation and

timing of cleavage are mechanisms that progress hand in hand, but one might act as a failsafe mechanism for the other when development is perturbed. It will be a future challenge to determine how these axes are initiated by the earliest events of embryogenesis and how they become transformed into the final body pattern.

Methods

Collection of eggs

Zygotes were collected from F₁ (C57BL/6x3BA) females induced to superovulate¹. Eggs were collected 14–15 h after hCG injection into PBS containing 200 IU ml⁻¹ hyaluronidase, dispersed and then transferred to FHM plus BSA medium as described¹.

Egg surface labelling and observations

Eggs were observed under an inverted (Leica) microscope using differential interference contrast (DIC) optics and micromanipulated with Leica micromanipulators using a De Fonbrune suction-force pump. Fluorescent (fluorescein-labelled) beads (3–4 µm diameter; Polysciences) were placed in FHM (Speciality Media, Laval, New Jersey) medium containing 300 µg ml⁻¹ phytohaemagglutinin for 30 min and then transferred to the chamber containing eggs in FHM plus bovine serum albumin. Individual beads were attached to the tip of a micropipette by suction, introduced through the zona pellucida and placed in contact with the membrane of the fertilization cone. Once the bead had adhered, the micropipette was withdrawn. Labelled eggs were transferred into KSOM (Speciality Media) medium and cultured as described² to the early blastocyst stage, when the ratio of the blastocoel and ICM is about 1/1. In some experiments a solution of 0.5 mg ml⁻¹ of rhodamine-labelled concanavalin A (Molecular Probes) in PBS was microinjected beneath the zona pellucida. The position of either of the markers on the embryo was scored using fluorescence and DIC microscopy. Live embryos were photographed using a cooled charge-coupled-device camera (Princeton Instruments) and DIC optics on a Nikon inverted microscope.

Received 9 August; accepted 2 November 2000.

- Weber, R. J., Pedersen, R. A., Wianny, F., Evans, M. J. & Zernicka-Goetz, M. Polarity of the mouse embryo is anticipated before implantation. *Development* **126**, 5591–5598 (1999).
- Ciemerych, M. A., Mesnard, D. & Zernicka-Goetz, M. Animal and vegetal poles of the mouse egg predict the polarity of the embryonic axis, yet are non-essential for development. *Development* **127**, 3467–3474 (2000).
- Gardner, R. L. The early blastocyst is bilaterally symmetrical and its axis of symmetry is aligned with the animal–vegetal axis of the zygote in the mouse. *Development* **124**, 289–301 (1997).
- Goldstein, B. & Hird, S. N. Specification of the anteroposterior axis in *Caenorhabditis elegans*. *Development* **122**, 1467–1474 (1996).
- Bowerman, B. in *Cell Lineage and Fate Determination* (ed. Moody, S. A.) 97–117 (Academic, San Diego, 1999).
- Gerhart, J. et al. Cortical rotation of the *Xenopus* egg: consequences for the anteroposterior pattern of embryonic dorsal development. *Development* **107**, 37–51 (1989).
- Graham, C. F. & Deussen, Z. A. Features of cell lineage in preimplantation mouse development. *J. Embryol. Exp. Morphol.* **48**, 53–72 (1978).
- Kelly, S. J., Mulnard, J. G. & Graham, C. F. Cell division and cell allocation in early mouse development. *J. Embryol. Exp. Morphol.* **48**, 37–51 (1978).
- Tarkowski, A. K. & Wroblewska, J. Development of blastomeres of mouse eggs isolated at the 4- and 8-cell stage. *J. Embryol. Exp. Morphol.* **18**, 155–180 (1967).
- Rossant, J. Postimplantation development of blastomeres isolated from 4- and 8-cell mouse eggs. *J. Embryol. Exp. Morphol.* **36**, 283–290 (1976).
- Balakier, H. & Pedersen, R. A. Allocation of cells to inner cell mass and trophectoderm lineages in preimplantation mouse embryos. *Dev. Biol.* **23**, 52–62 (1982).
- Surani, M. A. & Barton, S. C. Spatial distribution of blastomeres is dependent on cell division order and interactions in mouse morulae. *Dev. Biol.* **102**, 335–343 (1984).
- Garbutt, C. L., Johnson, M. H. & George, M. A. When and how does cell division order influence cell allocation to the inner cell mass of the mouse blastocyst? *Development* **100**, 325–332 (1987).
- Graham, C. F. & Lehtonen, E. Formation and consequences of cell patterns in preimplantation mouse development. *J. Embryol. Exp. Morphol.* **49**, 277–294 (1979).
- Garbutt, C. L., Chisholm, J. C. & Johnson, M. H. The establishment of the embryonic–abembryonic axis in the mouse embryo. *Development* **100**, 125–134 (1987).
- Copp, A. J. Interaction between inner cell mass and trophectoderm of the mouse blastocyst. II. The fate of the polar trophectoderm. *J. Embryol. Exp. Morphol.* **51**, 109–120 (1979).
- Cruz, Y. P. & Pedersen, R. A. Cell fate in the polar trophectoderm of mouse blastocysts as studied by microinjection of cell lineage tracers. *Dev. Biol.* **112**, 73–83 (1985).
- Gardner, R. L. Clonal analysis of growth of the polar trophectoderm in the mouse. *Hum. Reprod.* **11**, 1979–1984 (1996).
- Zernicka-Goetz, M. Fertile offspring derived from mammalian eggs lacking either animal or vegetal poles. *Development* **125**, 4803–4808 (1998).

Acknowledgements

We are grateful to D. Glover for invaluable suggestions and support; R. Pedersen and C. Graham for helpful discussions; and J. Gurdon for his critical comments. This work was supported by a Lister Institute of Preventive Medicine Fellowship, a Wellcome Trust Project Grant and a Royal Society Equipment Grant to M.Z.-G.

Correspondence and requests for materials should be addressed to M.Z.-G. (e-mail: mzg@mole.bio.cam.ac.uk).

Adherens junctions inhibit asymmetric division in the *Drosophila* epithelium

Bingwei Lu, Fabrice Roegiers, Lily Y. Jan & Yuh Nung Jan

Howard Hughes Medical Institute and Departments of Physiology and Biochemistry, University of California at San Francisco, San Francisco, California 94143-0725, USA

Asymmetric division is a fundamental mechanism for generating cellular diversity. In the central nervous system of *Drosophila*, neural progenitor cells called neuroblasts undergo asymmetric division along the apical–basal cellular axis^{1,2}. Neuroblasts originate from neuroepithelial cells, which are polarized along the apical–basal axis and divide symmetrically along the planar axis. The asymmetry of neuroblasts might arise from neuroblast-specific expression of the proteins required for asymmetric division. Alternatively, both neuroblasts and neuroepithelial cells could be capable of dividing asymmetrically, but in neuroepithelial cells other polarity cues might prevent asymmetric division. Here we show that by disrupting adherens junctions we can convert the symmetric epithelial division into asymmetric division. We further confirm that the adenomatous polyposis coli

(APC) tumour suppressor protein is recruited to adherens junctions³, and demonstrate that both APC and microtubule-associated EB1 homologues^{4–5} are required for the symmetric epithelial division along the planar axis. Our results indicate that neuroepithelial cells have all the necessary components to execute asymmetric division, but that this pathway is normally overridden by the planar polarity cue provided by adherens junctions.

Drosophila neuroblasts delaminate from a polarized epithelial layer in the ventral neuroectoderm and divide asymmetrically along the apical–basal axis to produce larger apical neuroblasts and smaller basal ganglion mother cells. Previous studies identified Inscuteable (Insc) as a central protein in organizing neuroblast division^{6,7}. Insc provides positional information that couples mitotic spindle orientation with the basal localization of cell-fate determinants such as Numb and Prospero together with their respective adaptor proteins Partner of Numb (Pon) and Miranda^{8–10}.

The apical localization of Insc involves both a Baz-dependent initiation step and a maintenance step that requires Baz and Partner of Inscuteable (Pins)^{11–15}. The expression of Baz and Pins in both neuroblasts and neuroepithelial cells suggests that these cells share certain apical–basal polarity information. Consistent with this notion is the observation that, when Pon is expressed ectopically in epithelial cells it is localized to the basal cortex, as in neuroblasts¹⁰. Unlike neuroblasts, however, epithelial cells divide symmetrically along the planar axis and segregate ectopic Pon equally between the two daughter cells. These observations raise further questions: do epithelial cells have the ability to couple spindle orientation with protein localization, and segregate proteins asymmetrically between two unequally sized daughter cells? If so, what prevents them from executing this asymmetric division?

To characterize epithelial division by monitoring it in live embryos, we used transgenic embryos expressing Pon and tau proteins fused with green fluorescent protein (GFP). During epithelial cell cycle, tau–GFP-labelled mitotic spindle is formed along the planar axis of the embryo, and Pon–GFP is initially uniformly associated with the cortex and then localized to a basal crescent. The mitotic spindle remains orientated along the planar

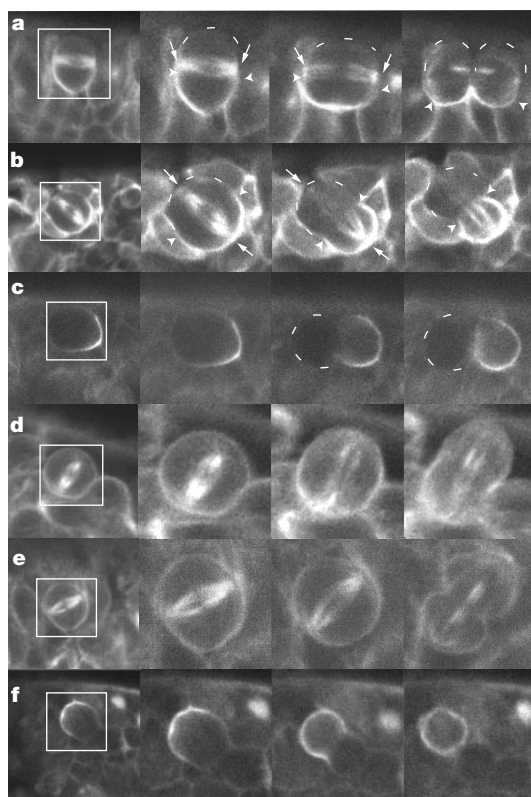


Figure 1 Epithelial divisions in wild-type and mutant embryos. **a**, Symmetric division of wild-type epithelial cells. **b**, **c**, **f**, Asymmetric epithelial divisions in *crb(RNAi)* (**b**), *UAS-Crb-intra* (**c**), and *crb^{11A22}/crb^{11A22}* (**f**) embryos. **d**, **e**, Symmetric epithelial division in *baz(RNAi)* (**d**) and *baz* and *crb* double RNAi (**e**) embryos. For each panel, the first image shows an epithelial cell under observation (square), followed by enlarged images of that cell during division. Arrowheads mark the Pon–GFP crescent, arrows mark tau–GFP-labelled spindles, and dotted lines mark the boundary of the Pon–GFP-negative daughter cells. Apical is up and basal down in all figures.

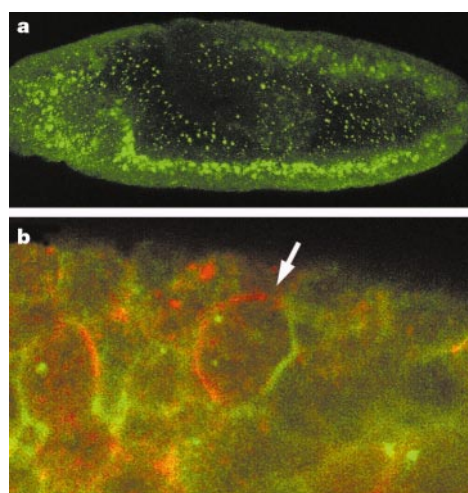


Figure 2 Mispositioning of Baz crescent and absence of ectopic neuronal marker expression in epithelial cells of *crb(RNAi)* embryos. **a**, *crb(RNAi)* embryos were allowed to develop to stages 10–12, and immunostained for Asense. Asense was expressed specifically in the developing central nervous system, and no ectopic expression in the epithelium was detected. **b**, Pon–GFP transgenic embryos injected with *crb* dsRNA were immunostained for Baz (red) and GFP (green). In the epithelial cell under observation (arrow), Baz is localized to a lateral crescent and Pon–GFP is localized as a crescent on the opposite side of the cortex.

axis throughout mitosis. After cytokinesis, the Pon–GFP crescent is bisected by the cleavage furrow and is equally distributed between two equally sized daughter cells (Fig. 1a). This *in vivo* analysis shows that the machinery for basal protein localization is intact in epithelial cells, but it is uncoupled from spindle orientation.

The uncoupling of spindle orientation with asymmetric protein localization in epithelial cells might be due to either a lack of such a coupling mechanism or the dominance of the coupling mechanism by yet another spindle-positioning mechanism. One of the hallmarks of epithelial cells is the adherens junction, which is composed of the cadherin–catenin complex and other associated proteins, is connected to the cytoskeleton, and is thought to be important in maintaining the planar organization of the epithelial monolayer. We therefore tested the possible role of adherens junction in orientating epithelial division. The formation of adherens junction requires genes such as *shotgun*, *crumbs* (*crb*) and *stardust*¹⁶. We used RNA interference (RNAi) to disrupt Crb function and analysed the effect on epithelial division^{17,18}.

We injected double-stranded (ds) *crb* RNA into transgenic embryos expressing Pon–GFP and tau–GFP. In about 70% ($n = 200$) of *crb*(RNAi) embryos, we observed that the organization of the ectodermal epithelium was disrupted, with epithelial cells losing their columnar shape, adopting rounded morphology, and becoming separated from each other. Live imaging of epithelial divisions in these embryos revealed that nearly all the epithelial cells show a tight coupling between the positioning of Pon–GFP crescents and the orientation of the mitotic spindle. Pon–GFP crescents were found at basal and lateral positions and less frequently at apical positions on the cell cortex, and one of the spindle poles was positioned underneath the Pon–GFP crescent.

After cytokinesis, Pon–GFP was segregated to one of the two similarly sized daughter cells (Fig. 1b). Asymmetric segregation of Pon–GFP to one of two similarly sized daughter cells was also observed in *crb* zygotic mutant embryos (Fig. 1f). Immunostaining

of *crb*(RNAi) embryos with antibodies against Asense, Prospero and Insc indicated that epithelial cells do not express these neuronal markers, suggesting that the ability of these cells to undergo asymmetric division is not a result of cell-fate change (Fig. 2a; and data not shown).

Overexpression of the membrane-bound cytoplasmic tail of Crb (Crb-intra) causes similar disorganization of the epithelium as seen in *crb* mutants^{19,20}. We therefore examined the effect of overexpressing Crb-intra on epithelial division. As observed in *crb*(RNAi) embryos, epithelial cells overexpressing Crb-intra showed coupling of the mitotic spindle with the Pon–GFP crescent and asymmetric segregation of Pon–GFP to one of the daughter cells (60%, $n = 100$) (Fig. 1c). Thus, when the formation of the adherens junction is disrupted, epithelial cells switch from a symmetric to an asymmetric division pattern.

In addition to its function in localizing Insc and regulating division axis in the neuroblasts, Baz is also required for the formation of adherens junction and the maintenance of epithelial polarity²¹. We next investigated the function of Baz in epithelial division. The *baz*(RNAi) embryos showed overall disruption of epithelium organization similar to that observed in *crb*(RNAi) embryos. Unlike in *crb*(RNAi) embryos, however, epithelial cells in *baz*(RNAi) embryos divided in a symmetric fashion, with Pon–GFP distributed uniformly around the cell cortex throughout mitosis and the mitotic spindle orientated in random directions. After cytokinesis, two equally sized daughter cells were produced and Pon–GFP was equally distributed between them (90%, $n = 300$) (Fig. 1d).

As observed previously¹³, however, daughter cell size asymmetry in neuroblast division was largely unaffected in *baz*(RNAi) embryos (data not shown). We also observed that in *crb*(RNAi) epithelial cells Baz can still be localized into a crescent but the crescent is mispositioned and that Pon–GFP is always localized to the opposite side of the Baz crescent (Fig. 2b). This suggests that, although

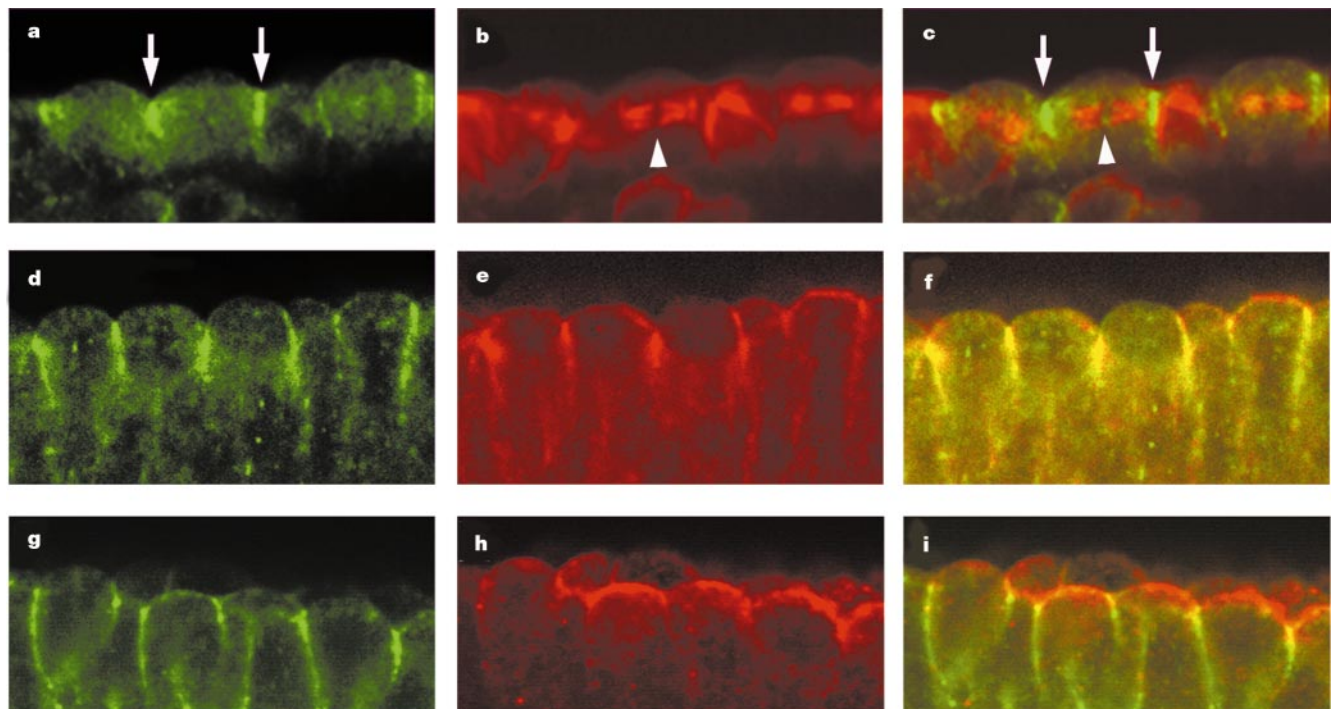


Figure 3 Localization of E-APC to the adherens junction and normal morphology of the adherens junction in epithelial cells in the PNR. **a–c**, E-APC (**a**) and tubulin (**b**) immunostaining. In dividing epithelial cells, the mitotic spindle (arrowhead) is orientated along the planar axis and E-APC is localized at cell–cell contacts (arrows). **d–f**, E-APC (**d**)

and E-cadherin (**e**) immunostaining. E-APC co-localizes with E-cadherin to the adherens junction. **g–i**, E-cadherin (**g**) and Insc (**h**) immunostaining. Epithelial cells within the PNR (Insc positive) exhibit normal adherens junction appearance compared with epithelial cells outside the PNR (compare **e** and **g**). **c**, **f** and **i** are merged images.

mispositioned, Baz is still functional in directing Pon–GFP localization in *crb(RNAi)* embryos. To test whether the coupling of Pon–GFP localization with spindle orientation observed in *crb(RNAi)* embryos is Baz dependent, we performed double RNAi by co-injecting a mixture of *baz* and *crb* dsRNAs. Epithelial divisions in the co-injected embryos looked similar to *baz* single-injected embryos, with Pon–GFP segregated equally between two equally sized daughter cells (80%, $n = 200$) (Fig. 1e). We therefore conclude that epithelial cells depend on Baz to couple spindle orientation with protein localization when the adherens junction is disrupted.

To investigate the molecular mechanism underlying the planar positioning of spindles by the adherens junction, we examined the function of proteins associated with the adherens junction. A ubiquitously expressed, epithelial-cell-enriched APC (E-APC) is localized to the adherens junction (Fig. 3a–f), and, in *shotgun* and *crb* mutants, this adherens junction localization of E-APC is disrupted^{3,22}. The human APC protein interacts with a microtubule-associated EB1 protein^{23,24}, and the yeast homologue of EB1 (Bim1), together with the cortical marker Kar9, has been implicated in a search-and-capture mechanism of spindle positioning^{4,5}. We therefore tested the function of E-APC in epithelial cell division.

In about 60% ($n = 300$) of *E-APC(RNAi)* embryos, we observed that the positioning of Pon–GFP crescent and orientation of mitotic spindle became tightly coupled during epithelial division. At cytokinesis, epithelial cells divided asymmetrically to produce two unequally sized daughter cells, and Pon–GFP was always segregated to the smaller daughter cell (Fig. 4a). The asymmetric segregation of Pon–GFP and the ability to undergo unequal cytokinesis all depend on Baz, because in *baz* and *E-APC* double RNAi embryos, Pon–GFP was equally segregated to two similarly sized daughter cells (data not shown). Therefore, in the absence of E-APC, epithelial cells divide asymmetrically in a Baz-dependent fashion. This suggests that adherens-junction-associated E-APC promotes spindle positioning along the planar axis and prevents the coupling of spindle positioning with asymmetric basal protein localization.

To test whether E-APC functions with EB1 to orientate the mitotic spindle, we performed RNAi on a closely related fly homologue of EB1 (*dEB1*). In *dEB1(RNAi)* embryos, the epithelial divisions were also asymmetric, producing two unequally sized daughter cells, with Pon–GFP segregated to the smaller cell (Fig. 4b). We observed that the penetrance of *dEB1(RNAi)* phenotype (~20%, $n = 300$) is lower than that of *E-APC(RNAi)*. As there is strong maternal contribution of *dEB1*, the low penetrance might be due to a perdurance of maternal *dEB1* protein. Alternatively, it might be due to functional compensation by two other distantly related EB1 homologues in the fly genome (data not shown). It has been noted that E-APC lacks the carboxy-terminal domain that is required for interaction with EB1 (ref. 3), and we did not detect a direct interaction between E-APC and EB1 in our *in vitro* binding assays. It therefore remains to be determined whether the two are functionally linked together *in vivo* through some cofactor(s), or whether E-APC functions mainly to maintain adherens junction integrity³ and EB1 interacts with other unidentified molecules to orientate spindles.

Our results indicate that two sets of polarity cues exist for spindle positioning in epithelial cells: a planar polarity cue mediated by the adherens junction and an apical–basal polarity cue regulated by Baz. The division pattern of wild-type epithelial cells suggests that the planar polarity cue is normally dominant over the apical–basal polarity cue (Fig. 4c). Epithelial cells within the procephalic neurogenic region (PNR) that express endogenous *Insc* or epithelial cells outside of the PNR that express ectopic *Insc* are known to orientate their mitotic spindle along the apical–basal axis during division⁷. This suggests that the dominance of planar polarity over apical–basal polarity can be overcome by the expression of *Insc*. The normal appearance of the adherens junction in epithelial cells in the PNR (Fig. 3g–i), together with the observation that these cells divide along the planar axis and maintain their normal monolayer

organization in *insc* mutant⁷, suggests that *Insc* functions by strengthening the apical–basal polarity instead of weakening the planar polarity through changing the behaviour of the adherens junction.

When neuroblasts delaminate from the epithelium layer, they undergo morphological changes from columnar to round shape, lose their contacts with the surrounding cells and thus the adherens junction structures. This situation may be reminiscent of epithelial cells in adherens-junction mutants in which the planar polarity cue is lost. In both cases, the Baz-mediated polarity pathway takes over. That one polarity cue can dominate over another cue in orientating division axis may have its precedents in other organisms. Budding yeast can divide in either an axial or a bipolar pattern. Mutations in genes such as *AXL1*, *BUD3*, *BUD4* and *BUD10/AXL2* result in loss of polarity cue for axial bud formation and the cells divide in a

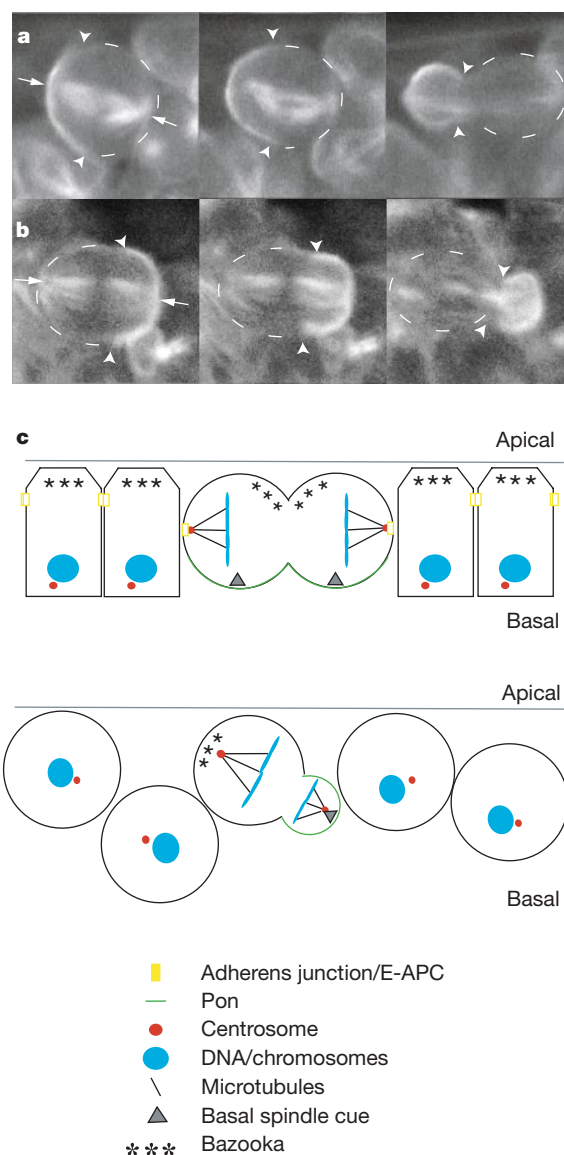


Figure 4 Epithelial divisions in *E-APC* and *dEB-1(RNAi)* mutant embryos. **a, b**, In an *E-APC(RNAi)* (**a**) and a *dEB-1(RNAi)* embryo (**b**), Pon–GFP forms a lateral crescent at mitosis, with the spindle coupled to it. At telophase, Pon–GFP is segregated asymmetrically to the smaller daughter cells. **c**, Model showing the competition between two polarity cues in controlling spindle positioning. In wild-type epithelial cells (top row), the planar polarity cue provided by adherens junction is dominant over the apical–basal polarity cue provided by Baz. When the planar polarity cue is lost, as in *E-APC*, *dEB-1*, and *crb* mutant embryos (bottom row), the Baz-mediated asymmetry cue takes over.

bipolar fashion^{25,26}. This suggests that axial and bipolar cues coexist and that the axial cue is normally dominant over the bipolar cue. During mammalian cortical neurogenesis, neural progenitors switch from early symmetric divisions to later asymmetric divisions^{27,28}. It will be interesting to determine whether similar mechanisms and molecules are used to control this division symmetry switch in mammals. Together with some recent studies²⁹, our results on E-APC highlights the importance of tumour suppressors in regulating not only cell growth but also polarity and asymmetric division. □

Methods

Fly stocks and genetics

We used the *UAS-GAL4* system to express ectopically *UAS-Pon-GFP* and *UAS-tau-GFP* (a gift from A. Brand) constructs in epithelial cells under the control of a maternal *V32A-GAL4* driver (a gift from D. St Johnston). For the RNA interference experiment, virgin females from a homozygous recombinant line of *V32A-GAL4* and *UAS-Pon-GFP* were crossed to males from a homozygous *UAS-tau-GFP* line, and embryos were collected for injection. For the overexpression of Crb intra, we crossed virgin females from a homozygous recombinant line of *V32A-GAL4* and *UAS-Pon-GFP* to males from a homozygous *UAS-Crb-intra* line (a gift from A. Wodarz). For characterization of epithelial cell division in *crumbs* mutant, *V32A-GAL4-UAS-Pon-GFP/+; crb^{11A22}/+* virgin females were crossed to males of the same genotype. Embryos produced from the above cross were aged to stages 9–10 and then processed for *in vivo* imaging study³⁰.

RNAi and *in vivo* imaging

Double-stranded RNAs were produced by *in vitro* transcription using polymerase chain reaction (PCR) products tagged at both ends with T7 RNA polymerase promoter sequences. The following PCR primers were used to generate the templates: dEB1, 5'-GGATCCTAATACGACTCACTATAGGAGGAGCCAGGAATCATTTAGTTCCTCCG; dEB1, 3'-GGATCCTA ATACGACTCACTATAGGAGGCGCTCTTTTCCAATCCCTCCAGG; E-APC, 5'-GGATCCTAATACGACTCACTATAGGAGGAGGAGTCGGAGGGTG AGCCGCCGGGG; E-APC, 3'-GGATCCTAATACGACTCACTATAGGAGGAGTGTCTGC AACTTGTAAATAATTAGCAGCTGGC; crb, 5'-GGATCCTAATACGACTCACTATAGG AGGT GGAATGGACAACGTAAGGAGCC; crb, 3'-GGATCCTAATACGACTCACTA TAGGAGGAGTACTCGGTATATATAGGATATAGG; baz, 5'-GGATCCTAATAC GACT ACTATAGGAGGAGCCGCCAGCAGCAACAGTTGGCAG; baz, 3'-GGATCCTAATAC GACTCACTATAGGAGGAGGACGTAGTGTCTCCATGGCCCTCGGC. Double-stranded RNAs were injected into *Pon-GFP* and *tau-GFP* transgenic embryos as described¹⁷. Aged embryos were subjected to *in vivo* imaging analysis³⁰.

Immunohistology

For immunostaining of wild-type and *crb(RNAi)* embryos, we processed an overnight collection of mixed-stage wild-type embryos, or *crb* dsRNA-injected embryos aged to stages 10–12 as described¹⁰. We used the following primary antibodies: rabbit anti-EAPC (a gift from M. Bienz), mouse anti-Tubulin (Sigma), rat anti-E-cadherin (a gift from T. Uemura), guinea pig anti-Asense, rabbit anti-Insc (a gift from B. Chia) and mouse anti-GFP (Molecular Probes), rat anti-Bazooka (a gift from A. Wodarz). Images were recorded on a confocal microscope and processed with Photoshop.

Received 20 October; accepted 20 November 2000.

1. Lu, B., Jan, L. Y. & Jan, Y. N. Control of cell divisions in the nervous system: symmetry and asymmetry. *Annu. Rev. Neurosci.* **23**, 531–556 (2000).
2. Jan, Y. N. & Jan, L. Y. Polarity in cell division: what frames thy fearful asymmetry? *Cell* **100**, 599–602 (2000).
3. Yu, X., Walzter, L. & Bienz, M. A new *Drosophila* APC homologue associated with adhesive zones of epithelial cells. *Nature Cell Biol.* **1**, 144–151 (1999).
4. Korinek, W. S., Copeland, M. J., Chaudhuri, A. & Chant, J. Molecular linkage underlying microtubule orientation toward cortical sites in yeast. *Science* **287**, 2257–2259 (2000).
5. Lee, L. *et al.* Positioning of the mitotic spindle by a cortical-microtubule capture mechanism. *Science* **287**, 2260–2262 (2000).
6. Kraut, R. & Campos-Ortega, J. A. *inscuteable*, a neural precursor gene of *Drosophila*, encodes a candidate for a cytoskeleton adapter protein. *Dev. Biol.* **174**, 65–81 (1996).
7. Kraut, R., Chia, W., Jan, L. Y., Jan, Y. N. & Knoblich, J. A. Role of *inscuteable* in orienting asymmetric cell division in *Drosophila*. *Nature* **383**, 50–55 (1996).
8. Ikeshima-Kataoka, H., Skeath, J. B., Nabeshima, Y., Doe, C. Q. & Matsuzaki, F. Miranda directs Prospero to a daughter cell during *Drosophila* asymmetric divisions. *Nature* **390**, 625–629 (1997).
9. Shen, C.-P., Jan, L. Y. & Jan, Y. N. Miranda is required for the asymmetric localization of Prospero during mitosis in *Drosophila*. *Cell* **90**, 449–458 (1997).
10. Lu, B., Rothenberg, M., Jan, L. Y. & Jan, Y. N. Partner of Numb colocalizes with Numb during mitosis and directs Numb asymmetric localization in *Drosophila* neural and muscle progenitors. *Cell* **95**, 225–235 (1998).
11. Kuchinke, U., Grawe, F. & Knust, E. Control of spindle orientation in *Drosophila* by the Par-3-related PDZ-domain protein Bazooka. *Curr. Biol.* **8**, 1357–1365 (1998).
12. Wodarz, A., Ramrath, A., Kuchinke, U. & Knust, E. Bazooka provides an apical cue for *inscuteable* localization in *Drosophila* neuroblasts. *Nature* **402**, 544–547 (1999).
13. Schober, M., Schaefer, M. & Knoblich, J. A. Bazooka recruits *inscuteable* to orient asymmetric cell divisions in *Drosophila* neuroblasts. *Nature* **402**, 548–551 (1999).

14. Yu, E., Morin, X., Cai, Y., Yang, X. & Chia, W. Analysis of Partner of *inscuteable*, a novel player of *Drosophila* asymmetric divisions, reveals two distinct steps in *inscuteable* apical localization. *Cell* **100**, 399–409 (2000).
15. Schaefer, M., Shevchenko, A., Shevchenko, A. & Knoblich, J. A. A protein complex containing *inscuteable* and the Gα-binding protein Pins orients asymmetric cell divisions in *Drosophila*. *Curr. Biol.* **10**, 353–362 (2000).
16. Tepass, U. Epithelial differentiation in *Drosophila*. *BioEssays* **19**, 673–682 (1997).
17. Kennerdell, J. R. & Carthew, R. W. Use of dsRNA-mediated genetic interference to demonstrate that *frizzled* and *frizzled 2* act in the *Wingless* pathway. *Cell* **95**, 1017–1026 (1998).
18. Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811 (1998).
19. Wodarz, A., Hinz, U., Engelbert, M. & Knust, E. Expression of *crumbs* confers apical character on plasma membrane domains of ectodermal epithelia. *Cell* **82**, 67–76 (1995).
20. Klebes, A. & Knust, E. A conserved motif in *Crums* is required for E-cadherin localisation and zonula adherens formation in *Drosophila*. *Curr. Biol.* **10**, 76–85 (2000).
21. Muller, H. A. & Wieschaus, E. *armadillo*, *bazooka*, and *stardust* are critical for early stages in formation of the zonula adherens and maintenance of the polarized blastoderm epithelium in *Drosophila*. *J. Cell Biol.* **134**, 149–163 (1996).
22. McCartney, B. M. *et al.* *Drosophila* APC2 is a cytoskeletally-associated protein that regulates wingless signaling in the embryonic epidermis. *J. Cell Biol.* **146**, 1303–1318 (1999).
23. Su, L. K. *et al.* APC binds to the novel protein EB1. *Cancer Res.* **55**, 2972–2977 (1995).
24. Nathke, I. S., Adams, C. L., Polakis, P., Sellin, J. H. & Nelson, W. J. The adenomatous polyposis coli tumor suppressor protein localizes to plasma membrane sites involved in active cell migration. *J. Cell Biol.* **134**, 165–179 (1996).
25. Chant, J. & Herskowitz, I. Genetic control of bud site selection in yeast by a set of gene products that constitute a morphogenetic pathway. *Cell* **65**, 1203–1212 (1991).
26. Roemer, T., Madden, K., Chang, J. & Snyder, M. Selection of axial growth sites in yeast requires Axl2p, a novel plasma membrane glycoprotein. *Genes Dev.* **10**, 777–793 (1996).
27. Caviness, V. S. J., Takahashi, T. & Nowakowski, R. S. Numbers, time and neocortical neurogenesis: A general developmental and evolutionary model. *Trends Neurosci.* **18**, 379–383 (1995).
28. Chenn, A. & McConnell, S. K. Cleavage orientation and the asymmetric inheritance of Notch1 immunoreactivity in mammalian neurogenesis. *Cell* **82**, 631–641 (1995).
29. Bilder, D., Li, M. & Perrimon, N. Cooperative regulation of cell polarity and growth by *Drosophila* tumor suppressors. *Science* **289**, 113–116 (2000).
30. Lu, B., Ackerman, L., Jan, L. Y. & Jan, Y. N. Modes of protein movement that lead to the asymmetric localization of Partner of Numb during neuroblast division in *Drosophila*. *Mol. Cell* **4**, 883–891 (1999).

Acknowledgements

We thank J. Chant for helpful suggestions; S. Guo and Y.-M. Chan for critically reading the manuscript and members of the Jan lab for stimulating discussions; A. Wodarz, A. Brand and D. St Johnston for providing fly strains; A. Wodarz, T. Uemura and B. Chia for antibodies; M. Bienz for antibodies and GST–E-APC constructs. This work was supported by a NIMH grant to the Silvo Conte Center for Neuroscience Research at UCSF, a National Research Service Award from NIH (B.L.) and a HSPF Fellowship (F.R.). Y.N.J. and L.Y.J. are Investigators of the Howard Hughes Medical Institute.

Correspondence and requests for materials should be addressed to Y.N.J. (e-mail: ynjan@itsa.ucsf.edu).

Complexes of MADS-box proteins are sufficient to convert leaves into floral organs

Takashi Honma & Koji Goto*

Institute for Chemical Research, Kyoto University, Uji, 611-0011, Japan

Genetic studies, using floral homeotic mutants, have led to the ABC model of flower development. This model proposes that the combinatorial action of three sets of genes, the A, B and C function genes, specify the four floral organs (sepals, petals, stamens and carpels) in the concentric floral whorls^{1,2}. However, attempts to convert vegetative organs into floral organs by altering the expression of ABC genes have been unsuccessful^{3–5}. Here we show that the class B proteins of *Arabidopsis*, PISTILLATA (PI) and APETALA3 (AP3), interact with APETALA1

* Present address: Research Institute for Biological Sciences, Kayo-cho, Jobo, Okayama, 716-1241, Japan.

(AP1, a class A protein) and SEPALLATA3 (SEP3, previously AGL9), and with AGAMOUS (AG, a class C protein) through SEP3. We also show that vegetative leaves of triply transgenic plants, 35S::PI;35S::AP3;35S::AP1 or 35S::PI;35S::AP3;35S::SEP3, are transformed into petaloid organs and that those of 35S::PI;35S::AP3;35S::SEP3;35S::AG are transformed into staminoid organs. Our findings indicate that the formation of ternary and quaternary complexes of ABC proteins may be the molecular basis of the ABC model, and that the flower-specific expression of SEP3 restricts the action of the ABC genes to the flower.

Most flowers consist of four types of floral organs in concentric whorls. The development of floral organs depends on the combinatorial action of genes, as proposed in the ABC model of flower development. This model proposes that the combinations of three classes of organ identity genes specify the four types of floral organs^{1,2}. That is, class A genes specify sepals in the first whorl, a combination of class A and B genes specify petals in the second whorl, class B and C genes specify stamens in the third whorl, and class C genes specify carpels in the fourth whorl. In *Arabidopsis*, *APETALA1* (AP1) and *APETALA2* (AP2) are class A genes, *PISTILLATA* (PI) and *APETALA3* (AP3) are class B genes, and *AGAMOUS* (AG) is classified as a class C gene. Molecular cloning of these genes revealed that AP1, PI, AP3 and AG encode members of the MADS family of transcription factors^{6–10}. Plant MADS proteins consist of four domains (Fig. 1d): the MADS (M) domain, a highly conserved DNA-binding domain; the I domain, an intervening region; the K domain, which is involved in protein–protein interactions; and the C domain¹¹. Homo- or heterodimers of MADS proteins recognize and bind the conserved DNA sequence, CC(A/T)₆GG, called the CArG box *in vitro*¹². The ABC model, however, suggests that different combinations of MADS proteins activate different groups of target genes in each whorls. How homologous MADS transcription factors are modulated to obtain

whorl-specific functions at the molecular level remains unknown. In addition, the ectopic expression of combinations of ABC genes does not result in the conversion of vegetative leaves into floral organs^{3–5}. This suggests that other unknown factors are required for floral organ identity.

Interactions with a ternary factor, either an unrelated protein or another MADS protein, may be responsible for the modulation of DNA-binding specificity and transcriptional activity¹³. PI and AP3 form a heterodimer to bind the CArG box *in vitro*^{7,12}. The AP3 promoter has three CArG boxes, which are bound by the PI–AP3 complex, and AP3 is autoregulated by PI–AP3 (refs 4, 14–16). In transgenic plants that express PI and AP3 constitutively (35S::PI;35S::AP3 plants), however, the AP3 promoter and *GUS* fusion gene (AP3::GUS) is expressed only in the floral organs^{4,16} (see Fig. 2a). When the transcriptional activation domain, VP16 (ref. 17), is fused with PI (35S::PI-VP16;35S::AP3), AP3::GUS is expressed throughout the plant (T.H. and K.G., unpublished data). These observations suggest that PI–AP3 requires a ternary factor that supplies a transcriptional activation domain and whose expression is flower-specific.

To identify the proposed ternary factor, we screened a flower complementary DNA library using the yeast two-hybrid system with both PI and AP3 as a bait. We found 340 positive colonies from 5.9×10^7 clones. We have sequenced 170 of the positive clones, and identified 19 clones of PI, 18 of SEP3, 15 of AP1 and 4 of ATA20 out of the clones that appeared more than twice. SEP3, AP1 or ATA20 can interact only when both PI and AP3 are expressed as a bait, and these interactions were confirmed by β -galactoside (β -gal) assay (Fig. 1a). SEP3 (previously called AGL9) is a member of MADS-box genes and is expressed in whorls 2–4 (ref. 18). As ATA20 is an anther-specific, putatively secreted protein¹⁹, we omitted ATA20 from further analysis. No other cofactor-like genes bearing transcriptional-activator domains were identified. These results suggest that PI–AP3 complex primarily interacts with AP1 and SEP3. In other words, the interactions among MADS proteins may be the principal protein–protein interactions of floral MADS proteins. We

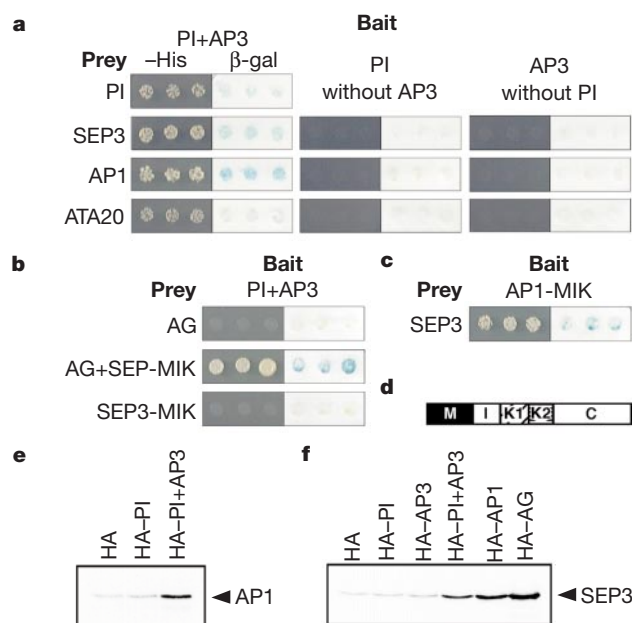


Figure 1 Interactions among MADS proteins in yeast and *in vitro*. **a**, Interactions between PI;LexA–AP3 and PI, SEP3, AP1 or ATA20 were confirmed by re-transformation on plates (–His) and a β -gal assay. No interaction occurred when either PI or AP3 was removed from the bait vector. **b**, Interactions between PI–AP3 and AG are mediated by SEP3. MIK indicates the deletion of the C domain. **c**, Interactions between SEP3 and AP1-MIK. **d**, Diagram of plant MADS proteins. The K domain can be divided into two regions (K1, K2) on the basis of α -helix formation²². **e**, **f**, Co-immunoprecipitation of MADS proteins. Results are representative of four independent experiments.

Table 1 Transactivation assay of MADS proteins

(a) In yeast			
MADS proteins	β -gal activity*		
	Full length	MIK†	K2C†
–‡	0.64 $\pm$ 0.03		
PI-VP16	148 $\pm$ 0.48		
PI	0.48 $\pm$ 0.02		
AP3	0.47 $\pm$ 0.01		
AP1	2.67 $\pm$ 0.14	0.49 $\pm$ 0.05	36.3 $\pm$ 0.45
SEP3	10.9 $\pm$ 0.63	0.46 $\pm$ 0.01	3.50 $\pm$ 0.17
AG	0.51 $\pm$ 0.01		
(b) In onion epidermal cells			
Effectors	Relative LUC/R-LUC activity§		
–¶	7.25 $\pm$ 0.88		
35S::PI	6.00 $\pm$ 1.11		
35S::PI-VP16	4.80 $\pm$ 1.01		
35S::AP3	4.85 $\pm$ 0.76		
35S::PI + 35S::AP3#	4.66 $\pm$ 0.70		
35S::PI-VP16 + 35S::AP3#	88.6 $\pm$ 12.8		
35S::AP1	55.1 $\pm$ 6.09		
35S::SEP3	142 $\pm$ 14.9		
35S::AG	4.80 $\pm$ 1.09		
35S::SEP1	63.1 $\pm$ 10.9		
35S::SEP2	13.0 $\pm$ 2.86		

* Mean ± s.e.m. × 10^{–1} Miller units. These data were calculated from five independent assays.

† Truncated protein deleting C domain (MIK) and MADS domain (K2C).

‡ GAL4-binding domain only (pAS2-1).

§ Transactivation activities were shown as arbitrary units of the LUC/R-LUC ratio from nine independent assays.

ll Mean ± s.e.m. × 10^{–2}. These data were calculated from five independent assays.

¶ Vector with 35S promoter and nos terminator was used.

Two plasmids were mixed.

further examined interactions between PI–AP3 and AG. AG does not interact with PI–AP3 directly (Fig. 1b), but AG and SEP3 interact in yeast²⁰. Yeast colonies survived only when all of *PI*, *AP3*, *AG* and *SEP3* were expressed (Fig. 1b). This result suggests that SEP3 mediates the interaction between PI–AP3 and AG.

We also examined the interactions between AP1 and SEP3, and found that these two proteins interact with each other (Fig. 1c). In these experiments, yeast, with full-length *AP1* or *SEP3* on the bait vector, survived without any prey. In contrast, when we used C-domain-deleted AP1 and SEP3 (*AP1-MIK* and *SEP3-MIK*), the yeast were not able to survive alone. This observation further suggests that AP1 and SEP3 have transcriptional activation domains²¹. The interactions of PI–AP3–AP1, PI–AP3–SEP3, AP1–SEP3 and AG–SEP3 were confirmed by immunoprecipitation experiments (Fig. 1e, f). As MADS proteins make a dimer to bind CArG boxes²², and the formation of a tetramer enhances the binding affinity to CArG-box repeats²³, PI–AP3–AP1–SEP3 and PI–AP3–SEP3–AG are the most likely complexes in the second and the third whorls, respectively.

How do these quaternary complexes alter the activity of MADS transcription factors? We thought that AP1 or SEP3 might add transcriptional-activator domains to PI–AP3 and AG, which do not possess them, and therefore that either AP1 or SEP3 might contain transcriptional-activator domains. To test this hypothesis, we measured the transcriptional activity of MADS proteins in yeast and plant cells (Table 1). Whereas PI, AP3 and AG do not have any detectable activity, AP1 and SEP3 possess moderate and strong activities, respectively, in both yeast and plant cells. In yeast, the transcriptional activity of both AP1 and SEP3 is abolished by the deletion of the C domain and is retained in the K2–C domain, the region that is sufficient for interactions with PI–AP3 (T.H. and K.G., unpublished data). In onion epidermal cells, *PI* can activate the reporter gene (*CArG::LUC*) only when both the activation domain was fused to PI (*PI-VP16*) and *AP3* was co-introduced, and other results agree with yeast experiments. These results suggest that AP1 and SEP3 have transcriptional-activator domains, which are localized primarily within their C domains, and are able to supply them to PI–AP3 or AG following complex formation. As the C domain is the most divergent region among the plant MADS proteins¹¹, it is feasible that some MADS proteins have transcriptional activity and others do not.

Interactions among the MADS proteins may also modulate the DNA-binding affinity and, thus, their target genes. Examples of these interactions are found in the yeast MADS protein, MCM1, and

in the *Antirrhinum* floral MADS proteins^{13,23,24}. We carried out two *in vivo* assays to examine the modulating ability of the complexes of the *Arabidopsis* MADS proteins.

First, we crossed the *AP3::GUS* gene into 35S::PI, 35S::AP3, 35S::AP1 and 35S::SEP3 plants, and into plants with combinations of these transgenes. *AP3::GUS* expression was observed in various tissues of both 35S::PI;35S::AP3;35S::AP1 and 35S::PI;35S::AP3;35S::SEP3 triply transgenic plants, whereas *AP3::GUS* was expressed only in the floral organs of 35S::AP1, 35S::SEP3, 35S::PI, 35S::AP3 or 35S::PI;35S::AP3 plants (Fig. 2; and data not shown). These results suggest that ternary complexes, PI–AP3–AP1 and PI–AP3–SEP3, are sufficient to activate the *AP3* promoter. As both AP1 and SEP3 form homodimers (data not shown), dimers probably provide the activation domain and then tetramers increase the DNA-binding affinity²³, although monomers of AP1 and SEP3 are sufficient to supply the activation domain to PI–AP3.

Second, we examined the phenotype of the above triply transgenic plants and 35S::PI;35S::AP3;35S::SEP3;35S::AG quadruply transgenic plants. These plants show remarkable phenotypes: vegetative leaves were transformed into floral organs (Fig. 3). Sixty per cent ($n = 55$) of 35S::SEP3 transgenic lines show a severe dwarf phenotype, curled leaves, early flowering and terminal flowers (Fig. 3a), with the remainder displaying an intermediate phenotype. 35S::PI;35S::AP3 plants exhibit curled leaves (Fig. 3b) and flowers in which the outer two whorls are petals and the inner two whorls are stamens⁴. These transgenic plants fail to convert vegetative leaves into floral organs, but the first true leaves of 35S::PI;35S::AP3;35S::SEP3 plants (the parental 35S::SEP3 displayed a severe phenotype) were converted into petaloid organs (Fig. 3c; compare g with e and f). These results suggest that PI–AP3–SEP3 is sufficient for the conversion of vegetative leaves into petaloid organs. As *ap1* mutants lack petals, AP1 is essential for petal identity. When SEP3 is over-expressed, however, its homodimer increases and then replaces AP1–SEP3 function. The observation that cauline leaves of 35S::PI;35S::AP3;35S::AP1 were also converted into petaloid organs (Fig. 3d, h) suggests that AP1 homodimer can function as AP1–SEP3. The functional redundancy of AP1 and SEP3 may reflect the fact that these two proteins are relatively closely related as they belong to AP1–SEP superclade^{25,26}. As 35S::SEP3 (severe);35S::AG plants are growth arrested and we were unable to obtain progeny, we used 35S::SEP3 (intermediate);35S::AG to construct 35S::PI;35S::AP3;35S::SEP3;35S::AG plants. Cauline leaves of this quadruply transgenic plant are converted into staminoid organs (Fig. 3i–o) and all floral organs are transformed into

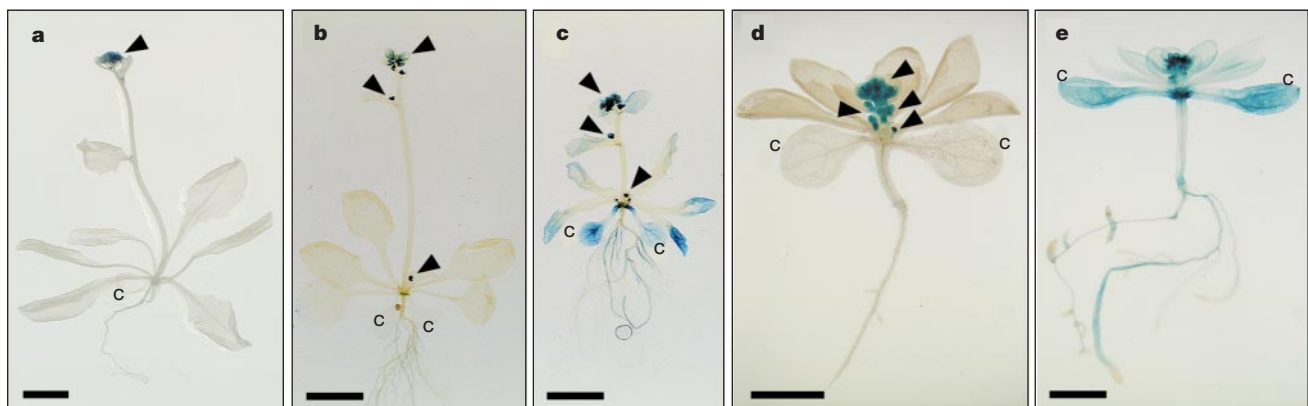


Figure 2 *AP3::GUS* expression in the transgenic plants. **a, b**, *AP3::GUS* in 35S::PI;35S::AP3 plants (**a**) and in 35S::AP1 plants (**b**). GUS activity was observed only in the flowers and floral buds (arrowheads). **c**, *AP3::GUS* in 35S::PI;35S::AP3;35S::AP1 triply transgenic plants. GUS activity was observed not only in floral organs (arrowheads), but also in roots, cotyledons, rosette and cauline leaves. **d**, *AP3::GUS* in a 3-week-old

35S::SEP3 plant. This line shows a severe phenotype. GUS expression is restricted to the floral organs (arrowheads). **e**, *AP3::GUS* in 17-day-old 35S::PI;35S::AP3;35S::SEP3 plants. GUS activity was observed in the whole plant. C, cotyledons. Scale bars, 5 mm (**a–c**); 1 mm (**d, e**).

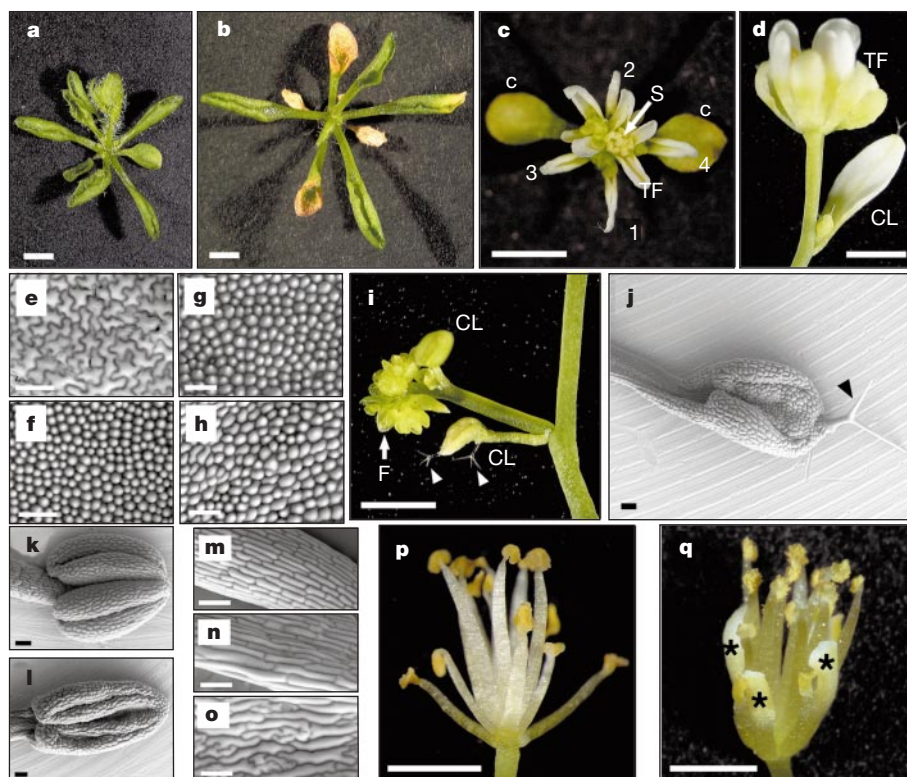


Figure 3 Phenotypes of triply and quadruply transgenic plants. **a**, Sixteen-day-old *35S::SEP3* plant showing a severe phenotype. **b**, Three-week-old *35S::PI;35S::AP3* plant displaying the curled leaf phenotype. **c**, Three-week-old *35S::PI;35S::AP3;35S::SEP3* plant. Cotyledons (C) are rather normal, but true leaves are transformed into petaloid organs. Numbers show the order of leaf development. S, stamens; TF, terminal flower. **d**, Three-week-old *35S::PI;35S::AP3;35S::AP1* plant. A cauline leaf (CL) is transformed into a petaloid organ. **e–h**, Cryo-scanning electron micrograph (cryo-SEM) of the adaxial surface of a *35S::SEP3* rosette leaf (**e**) which is similar to rosette and cauline leaves of the wild type; a wild-type petal (**f**); a petaloid *35S::PI;35S::AP3;35S::SEP3* rosette leaf (**g**); and a petaloid *35S::PI;35S::AP3;35S::AP1* cauline leaf (**h**). The epidermis of vegetative leaves consists of irregular 'jigsaw-puzzle-shaped' cells with interspersed stomata (**e**), whereas petal epidermis consists of conical ridged cells and lacks stomata (**f–h**).

i, Cauline leaves (CL) and lateral flowers (F) of *35S::PI;35S::AP3;35S::AG;35S::SEP3* quadruply transgenic plants. Not only all floral organs but also cauline leaves are transformed into stamens or staminoid organs. Arrowheads show branched trichomes. **j–o**, Cryo-SEM of staminoid cauline leaves of quadruply transgenic plants (**j**, **l**, **n**). For comparison, an anther (**k**), filament (**m**) and basal region of a cauline leaf (**o**) of wild-type are indicated. Wild-type cauline leaves lack a petiole and have irregularly shaped epidermal cells (**o**). Transformed cauline leaves of quadruply transgenic plants consist two distinct regions whose epidermal cells exhibit a morphology similar to that of wild-type anthers and filaments, respectively (**k–n**). **p**, **q**, Flowers of quadruply transgenic (**p**) and *35S::PI;35S::AP3;35S::AG* (**q**) plants. Asterisks show petaloid first whorl organs which are often incompletely converted into staminoid organs. Scale bars, 1 mm (**a–c**); 0.5 mm (**d**, **i**, **p**, **q**); 50 μ m (**e–h**, **j–o**).

stamens or staminoid organs (Fig. 3p). *35S::PI;35S::AP3;35S::AG* plants also have staminoid flowers, but first whorl organs, where *SEP3* is not expressed, often remain incompletely transformed (Fig. 3q), and they never display the conversion from vegetative leaves into floral organs. These results suggest that *PI–AP3–AG–SEP3* activity is sufficient for the conversion of leaves into staminoid organs.

We have shown that *SEP3* interacts with the *PI–AP3* complex and also serves as a scaffold between *PI–AP3* and *AG*. Ectopic expression of *PI–AP3–SEP3* and *PI–AP3–AP1* is sufficient to transform leaves into petaloid organs and that of *PI–AP3–SEP3–AG* is sufficient to convert cauline leaves into staminoid organs. These results indicate that floral organs can be formed independently of floral meristem, namely meristem identity and floral organ identity can be separated as suggested previously²⁷. A target gene of *PI–AP3* (*AP3::GUS*) is activated in non-floral organs when *SEP3* or *AP1* is expressed in addition to *PI–AP3*. Considering that *SEP3* and *AP1* can act as transcriptional activators, whereas *PI*, *AP3* and *AG* are not able to do so, *SEP3* and *AP1* can provide transactivation domains to the ternary and quaternary complexes. A loss-of-function allele of *SEP3* has been detected by T-DNA insertion screening⁵. The *sep3* mutant shows a subtle phenotype, but in combination with both *sep1* and *sep2*, mutants of *AGL2* and *AGL4*, it shows a similar phenotype to *bc* double mutants⁵. In onion epidermal cells, *SEP1* and *SEP2* show

moderate and weak transcriptional activity (Table 1), suggesting that these proteins can also supply activator domains to other MADS proteins. *SEP1* and *SEP2* genes are expressed in the floral whorls 1–4, and their redundant functions to *SEP3* suggest that *SEP* proteins confer completely functional activity on ABC proteins as transcription factors in the flower. *35S::SEP3*, *35S::SEP1* and *35S::SEP2* plants do not show homeotic change of floral organs (T.H. and K.G., unpublished data), but *35S::SEP3* together with *B*, *C* genes can convert leaves into floral organs. These results suggest that *SEP* genes provide the flower-specific activity to the function of the *B*, *C* genes by making complexes of their gene products. We propose that the ABC model should be amended to include *SEP* genes, which provide flower-specific activity. It is interesting to note that this functional divergence was acquired during the evolution of homologous MADS genes. □

Methods

Yeast two-hybrid screening and the β -gal assay

For the screening assay, both *LexA–AP3* (*AP3* fused with the LexA DNA-binding domain) and *PI* were expressed by the bait vector, pBTM116 (ref. 28), so that *PI* would be cloned as a positive control. The bait and the cDNA library on pACT2 were transformed in yeast L40 (ref. 28) simultaneously. We screened at 22 °C, the optimum temperature for *Arabidopsis* growth because the initial screening at 30 °C yielded no positive colonies. Interactions in yeast were confirmed by re-transforming three independent colonies on plates without histidine but with 2 mM 3-aminotriazole (–His), and by a colony-lift β -gal assay. To

examine the quaternary complex, *LexA*–*AP3* and *PI* were expressed on the bait vector, and *GAL4 AD-AG* and/or *SEP3-MIK* were expressed on the prey vector. When two genes were expressed on the same vector, they were both driven by ADH1 promoters. Amino-acid residues 1–167 and 1–171 were used for the truncated AP1-MIK and SEP3-MIK proteins, respectively. Other processes and the colony-lift β -gal assays were performed in accordance with the manufacturer's instructions (Clontech).

Immunoprecipitation

For immunoprecipitation experiments, radiolabelled AP1 or SEP3 were mixed with haemagglutinin (HA)-tagged proteins and precipitated with anti-HA antibody. Precipitated AP1 and SEP3 were separated by SDS–PAGE and detected by radio-imaging analyser, BAS2000 (Fujifilm). Other procedures were done as described^{7,12}.

Transactivation assay

For yeast, MADS proteins cDNAs were fused in-frame to GAL4 DNA-binding domain on pAS2-1 (Clontech) and transformed into the yeast strain YRG-2 (*UAS::lacZ*, Stratagene). AP1-K2C (residues 125–256) and SEP3-K2C (128–257) were used as truncated MADS proteins. Yeast cells were grown at 22 °C overnight, and the β -gal activity was assayed at 30 °C using *o*-nitrophenyl- β -D-galactopyranoside.

For onion epidermal cells, 35S promoter-driven MADS cDNAs that express native MADS proteins (effector) and *CaRG::LUC* (reporter) were co-transfected into onion epidermal cells by using a particle delivery system (Bio-Rad). *CaRG::LUC* has seven repeats of MADS protein binding consensus sequence²⁹, 5'-GGGGTGGCTTCTCTTTTGG TAAATTTTGGATCC-3' (*CaRG* box is underlined), upstream of the 35S minimal promoter (–30). 35S::*Renilla* luciferase (RLUC) was used for the internal control. LUC assays were conducted using Dual-luciferase reporter system (Promega). Other procedures were done as described³⁰.

Plant material

Arabidopsis Columbia ecotype was used for *Agrobacterium*-mediated vacuum transformation³¹. Plant crossing was carried out by manual cross-pollination. The presence of the transgenes was confirmed by PCR. *AP3::GUS* plants have a 600-base-pair region of the *AP3* promoter¹⁶. Staining for GUS activity was done as described¹⁶.

Cryo-scanning electron micrograph

We used a Hitachi S-3500N scanning electron microscope equipped with a cryo-stage. For observation and photography, the stage was chilled at –20 °C and the natural scanning electron microscopy (SEM) mode (70 Pa) was used with a 25-kV accelerating voltage.

Received 2 October; accepted 6 November 2000.

- Coen, E. S. & Meyerowitz, E. M. The war of the whorls: genetic interactions controlling flower development. *Nature* **353**, 31–37 (1991).
- Bowman, J. L., Smyth, D. R. & Meyerowitz, E. M. Genetic interactions among floral homeotic genes of *Arabidopsis*. *Development* **112**, 1–20 (1991).
- Mizukami, Y. & Ma, H. Ectopic expression of the floral homeotic gene *AGAMOUS* in transgenic *Arabidopsis* plants alters floral organ identity. *Cell* **71**, 119–131 (1992).
- Krizek, B. A. & Meyerowitz, E. M. The *Arabidopsis* homeotic genes *APETALA3* and *PISTILLATA* are sufficient to provide the B class organ identity function. *Development* **122**, 11–22 (1996).
- Pelaz, S., Ditta, G. S., Baumann, E., Wisman, E. & Yanofsky, M. F. B and C floral organ identity functions require *SEPALLATA* MADS-box genes. *Nature* **405**, 200–203 (2000).
- Mandel, M. A., Gustafson-Brown, C., Savidge, B. & Yanofsky, M. F. Molecular characterization of the *Arabidopsis* floral homeotic gene *APETALA1*. *Nature* **360**, 273–277 (1992).
- Goto, K. & Meyerowitz, E. M. Function and regulation of the *Arabidopsis* floral homeotic gene *PISTILLATA*. *Genes Dev.* **8**, 1548–1560 (1994).
- Jack, T., Brockman, L. L. & Meyerowitz, E. M. The homeotic gene *APETALA3* of *Arabidopsis thaliana* encodes a MADS box and is expressed in petals and stamens. *Cell* **68**, 683–697 (1992).
- Yanofsky, M. F. *et al.* The protein encoded by the *Arabidopsis* homeotic gene *agamous* resembles transcription factors. *Nature* **346**, 35–39 (1990).
- Schwarz-Sommer, Z., Huijser, P., Nacken, W., Saedler, H. & Sommer, H. Genetic control of flower development: homeotic genes in *Antirrhinum majus*. *Science* **250**, 931–936 (1990).
- Ma, H., Yanofsky, M. F. & Meyerowitz, E. M. *AGL1*–*AGL6*, an *Arabidopsis* gene family with similarity to floral homeotic and transcription factor genes. *Genes Dev.* **5**, 484–495 (1991).
- Riechmann, J. L., Krizek, B. A. & Meyerowitz, E. M. Dimerization specificity of *Arabidopsis* MADS domain homeotic proteins *APETALA1*, *APETALA3*, *PISTILLATA*, and *AGAMOUS*. *Proc. Natl Acad. Sci. USA* **93**, 4793–4798 (1996).
- Herskowitz, I. A regulatory hierarchy for cell specialization in yeast. *Nature* **342**, 749–757 (1989).
- Tilly, J. J., Allen, D. W. & Jack, T. The *CaRG* boxes in the promoter of the *Arabidopsis* floral organ identity gene *APETALA3* mediate diverse regulatory effects. *Development* **125**, 1647–1657 (1998).
- Hill, T. A., Day, C. D., Zondlo, S. C., Thackeray, A. G. & Irish, V. F. Discrete spatial and temporal cis-acting elements regulate transcription of the *Arabidopsis* floral homeotic gene *APETALA3*. *Development* **125**, 1711–1721 (1998).
- Honma, T. & Goto, K. The *Arabidopsis* floral homeotic gene *PISTILLATA* is regulated by discrete cis-elements responsive to induction and maintenance signals. *Development* **127**, 2021–2030 (2000).
- Sadowski, I., Ma, J., Triesenberg, S. & Ptashne, M. *GAL4*-VP16 is an unusually potent transcriptional activator. *Nature* **335**, 563–564 (1988).
- Mandel, M. A. & Yanofsky, M. F. The *Arabidopsis* *AGL9* MADS box gene is expressed in young flower primordia. *Sex Plant Reprod.* **11**, 22–28 (1998).
- Rubinelli, P., Hu, Y. & Ma, H. Identification, sequence analysis and expression studies of novel anther-specific genes of *Arabidopsis thaliana*. *Plant Mol. Biol.* **37**, 607–619 (1998).

- Fan, H. -Y., Hu, Y., Tudor, M. & Ma, H. Specific interactions between the K domains of AG and AGLs, members of the MADS domain family of DNA binding proteins. *Plant J.* **12**, 999–1010 (1997).
- Choi, S. *et al.* Analysis of the C-terminal region of *Arabidopsis thaliana* APETALA1 as a transcription activation domain. *Plant Mol. Biol.* **40**, 419–429 (1999).
- Riechmann, J. L. & Meyerowitz, E. M. MADS domain proteins in plant development. *J. Biol. Chem.* **378**, 1079–1101 (1997).
- Egea-Cortines, M., Saedler, H. & Sommer, H. Ternary complex formation between the MADS-box proteins *SQUAMOSA*, *DEFICIENS* and *GLOBOSA* is involved in the control of floral architecture in *Antirrhinum majus*. *EMBO J.* **18**, 5370–5379 (1999).
- Davies, B., Egea-Cortines, M., de Andrade Silva, E., Saedler, H. & Sommer, H. Multiple interactions amongst floral homeotic MADS box proteins. *EMBO J.* **15**, 4330–4343 (1996).
- Rounsley, S. D., Ditta, G. S. & Yanofsky, M. F. Diverse roles for MADS box genes in *Arabidopsis* development. *Plant Cell* **7**, 1259–1269 (1995).
- Smyth, D. A reverse trend—MADS functions revealed. *Trends Plant Sci.* **5**, 315–317 (2000).
- Parcy, F., Nilsson, O., Busch, M. A., Lee, I. & Weigel, D. A genetic framework for floral patterning. *Nature* **395**, 561–566 (1998).
- Bartel, P. L., Chien, C., Sternglanz, R. & Fields, S. in *Cellular Interactions in Development: a Practical Approach*. (ed. Hartley, D. A.) 153–179 (IRL Press, Oxford, 1993).
- Shiraishi, H., Okada, K. & Shimura, Y. Nucleotide sequences recognized by the *AGAMOUS* MADS domain of *Arabidopsis thaliana* in vitro. *Plant J.* **4**, 385–398 (1993).
- Pan, S., Sehne, P. C., Ferl, R. J. & Gurley, W. B. Specific interactions with TBP and TFIIB in vitro suggest that 14-3-3 proteins may participate in the regulation of transcription when part of a DNA binding complex. *Plant Cell* **11**, 1591–1602 (1999).
- Bechtold, N., Ellis, J. & Pelletier, G. In *planta Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis* plants. *C. R. Acad. Sci. Paris* **316**, 1194–1199 (1993).

Acknowledgements

We are grateful to M. Yanofsky for communicating data before publication, and to D. Weigel for providing the cDNA library. We also thank J. Bowman, T. Ito and H. Tsukaya for critical reading of the manuscript. This work was supported by grants from the Monbusho and JSPS.

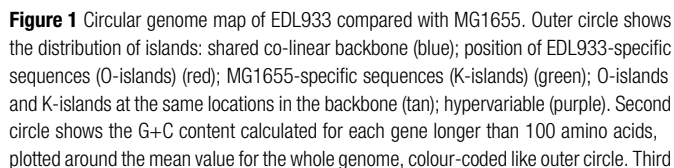
Correspondence and requests for materials should be addressed to K.G. (e-mail: kgoto@v004.vaio.ne.jp).

Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7

Nicole T. Perna[†]*, Guy Plunkett III[‡], Valerie Burland[‡], Bob Mau[‡], Jeremy D. Glasner[‡], Debra J. Rose[‡], George F. Mayhew[‡], Peter S. Evans[‡], Jason Gregor[‡], Heather A. Kirkpatrick[‡], György Pósfai[§], Jeremiah Hackett[‡], Sara Klink[‡], Adam Boutin[‡], Ying Shao[‡], Leslie Miller[‡], Erik J. Grobeck[‡], N. Wayne Davis[‡], Alex Lim^{||}, Eileen T. Dimalanta^{||}, Konstantinos D. Potamou^{||}, Jennifer Apodaca[‡]||, Thomas S. Anantharaman[¶], Jieyi Lin[#], Galex Yen^{*}, David C. Schwartz^{*‡}||, Rodney A. Welch[‡] & Frederick R. Blattner^{*‡}

* Genome Center of Wisconsin, † Department of Animal Health and Biomedical Sciences, ‡ Laboratory of Genetics, || Department of Chemistry, ¶ Department of Biostatistics, and [‡] Department of Medical Microbiology and Immunology, University of Wisconsin, Madison, Wisconsin 53706, USA
[§] Institute of Biochemistry, Biological Research Center, H-6701 Szeged, Hungary
[#] Cereon Genomics, LLC, 45 Sidney Street, Cambridge, Massachusetts 02139, USA

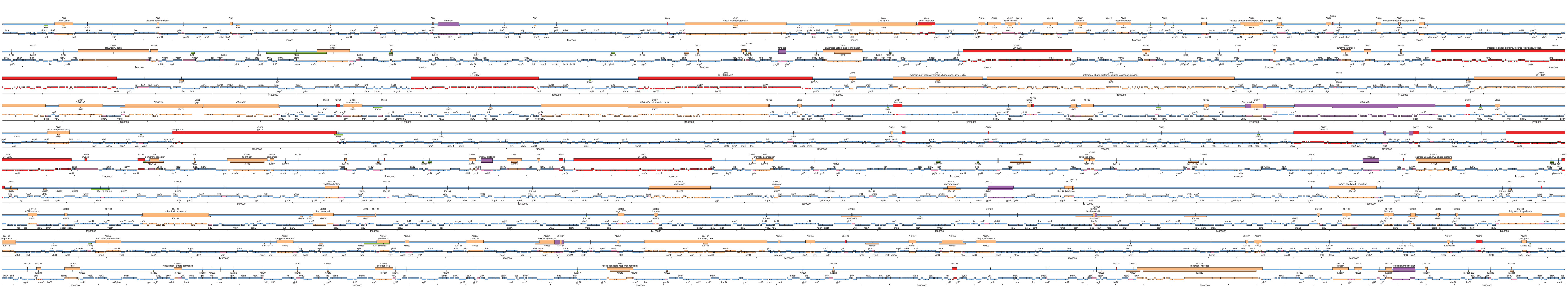
The bacterium *Escherichia coli* O157:H7 is a worldwide threat to public health and has been implicated in many outbreaks of haemorrhagic colitis, some of which included fatalities caused by haemolytic uraemic syndrome^{1,2}. Close to 75,000 cases of O157:H7 infection are now estimated to occur annually in the United States³. The severity of disease, the lack of effective treatment and the potential for large-scale outbreaks from contaminated food supplies have propelled intensive research on the pathogenesis and detection of *E. coli* O157:H7 (ref. 4). Here we have sequenced the genome of *E. coli* O157:H7 to identify candidate genes responsible for pathogenesis, to develop better methods of strain detection and to advance our understanding of



Escherichia coli O157:H7 was first associated with human disease after a multi-state outbreak in 1982 involving contaminated hamburgers¹. The strain EDL933 that we sequenced was isolated from Michigan ground beef linked to this incident, and has been studied as a reference strain for O157:H7. Figures 1 and 2 show the gene content and organization of the EDL933 genome, and compare it with the chromosome of the K-12 laboratory strain MG1655 (ref. 5). These strains last shared a common ancestor about 4.5 million years ago⁶. The two *E. coli* genomes revealed an unexpectedly complex segmented relationship, even in a preliminary examination⁷. They share a common 'backbone' sequence which is co-linear except for one 422-kilobase (kb) inversion spanning the replication terminus (Fig. 1). Homology is punctuated by hundreds of islands of apparently introgressed DNA—numbered and designated 'K-islands' (KI) or 'O-islands' (OI) in Fig. 2, where K-islands are DNA segments present in MG1655 but not in EDL933, and O-islands are unique segments present in EDL933.

The backbone comprises 4.1 megabases (Mb), which are clearly homologous between the two *E. coli* genomes. O-islands total 1.34 Mb of DNA and K-islands total 0.53 Mb. These lineage-specific segments are found throughout both genomes in clusters of up to 88 kb. There are 177 O-islands and 234 K-islands greater than 50 bp in length. Histograms (Fig. 3) show more intermediate and large islands in EDL933 than in MG1655. Only 14.7% (26/177) of the O-islands correspond entirely to intergene regions. The two largest are identical copies of a 106-gene island, both in the same orientation and adjacent to genes encoding identical transfer RNAs.

530



Labelling lineage-specific segments 'islands' is an extension of the term 'pathogenicity island' now in common, albeit varied, use. The original term arose from observations that virulence determinants are often clustered in large genomic segments showing hallmarks of horizontal transfer⁸. However, we found K- and O-islands of all sizes with no obvious association with pathogenicity; conversely, genes probably associated with virulence are not limited to the largest islands.

Roughly 26% of the EDL933 genes (1,387/5,416) lie completely within O-islands. In 189 cases, backbone-island junctions are within predicted genes. We classified the EDL933 genes into the functional groups reported for the MG1655 genome³ and this is included in the annotation. Of the O-island genes, 40% (561) can be assigned a function. Another 338 EDL933 genes marked as unknowns lie within phage-related clusters and are probably remnants of phage genomes. About 33% (59/177) of the O-islands contain only genes of unknown function. Many classifiable proteins are related to known virulence-associated proteins from other *E. coli* strains or related enterobacteria.

Nine large O-islands (>15 kb) encode putative virulence factors: a macrophage toxin and ClpB-like chaperone (OI#7); a RTX-toxin-like exoprotein and transport system (OI#28); two urease gene clusters (OI#43 and #48); an adhesin and polyketide or fatty-acid biosynthesis system (OI#47); a type III secretion system and secreted proteins similar to the *Salmonella-Shigella inv-spa* host-cell invasion genes (OI#115); two toxins and a PagC-like virulence factor (OI#122); a fatty-acid biosynthesis system (OI#138); and the previously described locus of enterocyte effacement (OI#148)⁹. Among the large islands, four include a P4-family integrase and are directly adjacent to tRNAs (OI#43-*serW*, #48-*serX*, #122-*pheV* and #148-*selC*). Only the locus of enterocyte effacement and two of the lambdoid phages (see below) have as yet been experimentally associated with virulence in animal models.

Smaller islands that may be involved in virulence contain fimbrial biosynthesis systems, iron uptake and utilization clusters, and putative non-fimbrial adhesins. Many clusters have no obvious role in virulence, but may confer strain-specific abilities to survive in different niches. Examples include candidates for transporting diverse carbohydrates, antibiotic efflux, aromatic compound degradation, tellurite resistance and glutamate fermentation. Not all islands are expected to be adaptive. Some may represent neutral variation between strains. Still others may be deleterious but either have not yet been eliminated by selection or cannot be eliminated because of linkage constraints.

We identified 18 multigenic regions of the EDL933 chromosome related to known bacteriophages. Only one, the Stx2 Shiga toxin-converting phage BP-933W, is known to be capable of lytic growth and production of infectious particles¹⁰. We named the other EDL933 prophages cryptic prophage (CP) to indicate that they probably lack a full complement of functional phage genes. They vary in size from 7.5 kb (CP-933L) to 61.6 kb (BP-933W) and consist of a mosaic of segments similar to various bacteriophages, recalling the 'modular' phage genome hypothesis¹¹. The two remaining physical gaps in the genome sequence correspond to prophage-related regions, and resolution of the sequence is complicated by extensive similarities to other prophage within this genome. The gap sizes and positions (4 kb and 54 kb) were determined from optical restriction maps. With only one exception, the EDL933 prophages and the eight cryptic prophages of MG1655 are all lineage-specific. Prophage Rac (MG1655) and CP-933R are similarly located in the backbone, and are sufficiently related to suggest a common prophage ancestor at the time that the strains diverged.

Subunits Stx1A and Stx1B of the second Shiga toxin of EDL933 (ref. 12) are encoded in the newly identified CP-933V. The position of the *stx1* genes in a putative Q antiterminator-dependent transcript is analogous to the placement of the *stx2* genes in BP-933W,

although there are no tRNAs adjacent to *stx1AB*. Genes in this position should be expressed maximally during lytic growth. The relationship between Stx toxin expression and phage induction is important, because treatment of O157:H7 with macrolide and quinolone antibiotics increase expression of the toxins^{13,14}. Clinical decisions regarding drug therapy are complicated by strain-specific variation in this response¹⁵, and reports in the literature (for example, refs 6 and 12) taken together suggest that the Stx phage status is variable among O157:H7 strains. Given the potential for recombination among the prophage reported here, this does not seem surprising. In addition, the *stx* locus in *Shigella* is known to lie within a cryptic prophage, inserted at a site different from either *stx* phage of EDL933 (ref. 16).

The MG1655 genome contains 528 genes (528/4,405 = 12%) not found in EDL933. About 57% (303) of these were classified into known functional groups and include genes, such as for ferric citrate utilization, that would suggest a role in virulence if identified in a pathogen. It is unclear whether these are remnants of a recent pathogenic ancestor, steps along a path leading to evolution of a new pathogen, indicators that K-12 strains may be pathogenic for non-human hosts, or completely unrelated to pathogenicity. There are 106 examples of O-islands and K-islands present at the same locations relative to the conserved chromosomal backbone. The two replichores in each strain are nearly equal in length despite the large number of insertion/deletion events necessary to generate the observed segmented structure between strains. Only a subset of islands is associated with elements likely to be autonomously mobile.

Each island might be ancestral and lost from the reciprocal genome; however, atypical base composition suggests that most islands are horizontal transfers of relatively recent origin from a donor species with a different intrinsic base composition. Restricting analysis to the 108 O-islands greater than 1 kb, 94% (101/108) are significantly different ($\chi^2 > 7.815$, $P < 0.05$) from the average base composition of shared backbone regions in the same replicore. The percentage drops very little with a Bonferroni correction for multiple tests (91/108; $\chi^2 > 17.892$, $P < 0.05$). Similar results are obtained for analysis of the third-codon position composition (Fig. 1). Still more islands may have originated as horizontal transfers but have been resident in genomes with a spectrum of mutation similar enough to *E. coli* to have obtained equilibrium

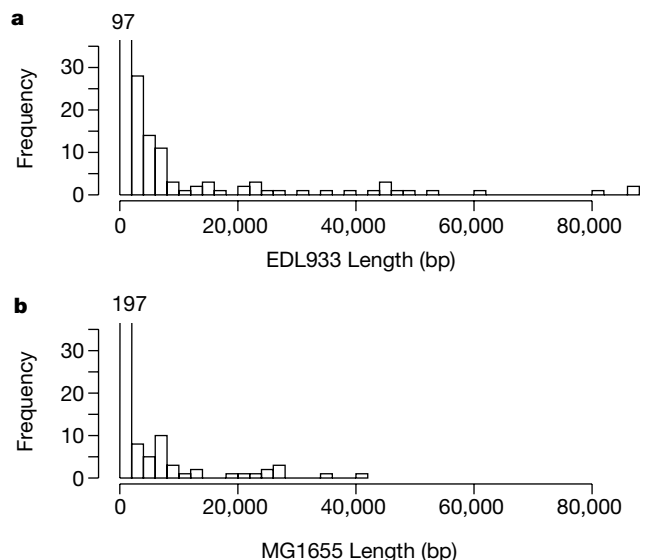


Figure 3 Histograms of lineage-specific segment lengths. **a**, EDL933; **b**, MG1655. The frequencies for the smallest length class are truncated to emphasize the distribution of longer clusters.

nucleotide frequencies or at least obscure statistical significance¹⁷. Still other gene clusters may be horizontal transfers that predate the divergence of MG1655 and EDL933.

Single nucleotide polymorphisms (75,168 differences) are distributed throughout the homologous backbone. There are 3,574 protein-coding genes encoded in backbone, and the average nucleotide identity for orthologous genes is 98.4%. Many orthologues (3,181/3,574 = 89%) are of equal length in the two genomes, but only 25% (911) encode identical proteins. Table 1 shows the number of each type of polymorphism observed by codon position. As expected, most differences are synonymous changes at third-codon positions. Multiple mutations at the same site should be infrequent at this low level of divergence. Thus the co-occurrence matrix provides insight into the substitution pattern, despite uncertainty of the ancestral state. The overall ratio of transitions to transversions is close to 3:1. A bias towards a greater number of T→C than A→G transitions on the coding strand previously attributed to transcription-coupled repair is evident¹⁸. An additional bias was observed at third-codon positions. Thymidines are more frequently involved in transversions than cytosines, and G→T are the most frequent transversions for the coding strand. The reciprocal polymorphisms, C→A, are not over-represented. This bias is consistent for genes on both the leading and lagging strands (data not shown) and is therefore not related to asymmetries in the replication process. One possible explanation is transcription-coupled repair of damage associated with oxidative stress. Oxidized products of guanine (2,2,4-triaminooxazolone and 7,8-dihydro-8-oxoguanine) lead to G→T transversions by mispairing with A, and two DNA glycosylases (MutY and Fpg) are responsible for mismatch resolution¹⁹. Preferential repair of these lesions on the transcribed strand has been observed in humans²⁰, and a similar mechanism could account for the observed transversion bias on the coding strand in *E. coli*.

Some chromosomal regions are more divergent ('hypervariable') than the average homologous segment but encode a comparable set of proteins at the same relative chromosomal position. In the most extreme case (YadC), the MG1655 and EDL933 proteins exhibit only 34% identity. Four such loci encode known or putative fimbrial

biosynthesis operons. Another encodes a restriction/modification system. Elevated divergence has been associated with positive selection at both these types of loci and among proteins that interact directly with the host^{9,21,22}. Alternatively, hypervariable genes may result from locally elevated mutation rates or differential paralogue retention from an ancient tandem duplication.

Comparison of our observations with other genome-scale analyses of closely related strains or species supports the idea that enterobacterial genomes are particularly subject to recombinational evolution. Two *Helicobacter pylori* strains exhibited only 6–7% differential coding capacity despite showing less identity among orthologues (92.6%) than observed among these *E. coli*. Furthermore, almost half of the lineage-specific *Helicobacter* genes are clustered in a single region referred to as the plasticity zone²³. Analyses of four *Chlamydia* genomes with orthologues that differ by as much as 19.5% show little evidence of horizontal transfer, and this is attributed to the inherent isolation of an obligate intracellular parasite²⁴. Most lineage-specific genes are expansions of paralogous gene families. As in *Helicobacter*, many of the *Chlamydia* lineage-specific elements are clustered in a plasticity zone. Continuing genome projects will elucidate the generality of observations made from these comparisons of closely related organisms.

Together, our findings reveal a surprising level of diversity between two members of the species *E. coli*. Most differences in overall gene content are attributable to horizontal transfer, and offer a wealth of candidate genes that may be involved in pathogenesis. Base substitution has introduced variation into most gene products even among conserved regions of the two strains. Many of these differences can be exploited for development of highly sensitive diagnostic tools; but diagnostic utility will require a clearer understanding of the distribution of genetic elements in *E. coli* species as a whole. An independently isolated O157:H7 strain showed differences from EDL933 by restriction mapping²⁵. Additional genome sequence data from other *E. coli* strains as well as functional characterization of gene products is necessary before the complex relationship between *E. coli* genotypes and phenotypes can be understood. Showing that disease-related traits are associated with predicted genes will require many areas of study including extensive testing in animal models that mimic symptoms of human infections, but the genome sequence offers a unique resource to help meet the challenge. □

Table 1 Frequency of each type of single nucleotide polymorphism by codon position for 3,181 genes of equal length in EDL933 and MG1655

First-codon position		Base in MG1655				EDL933 totals
Base in EDL933		G	A	T	C	
G	–	865	154	137	1,156	
A	924	–	129	286	1,339	
T	170	143	–	1,260	1,573	
C	124	284	1,156	–	1,564	
MG1655 totals	1,218	1,292	1,439	1,683	5,632	
Second-codon position		Base in MG1655				EDL933 totals
Base in EDL933		G	A	T	C	
G	–	410	66	118	594	
A	393	–	159	166	718	
T	67	147	–	464	678	
C	107	176	394	–	677	
MG1655 totals	567	733	619	748	2,667	
Third-codon position		Base in MG1655				EDL933 totals
Base in EDL933		G	A	T	C	
G	–	6,021	1,562	1,024	8,607	
A	6,107	–	1,242	1,124	8,473	
T	1,619	1,228	–	8,538	11,385	
C	1,049	1,010	8,307	–	10,366	
MG1655 totals	8,775	8,259	11,111	10,686	38,831	

Methods

Clones and sequencing

EDL933 was kindly provided by C. Kaspar, who obtained it from the American Type Culture Collection (ATCC 43895). The sequenced isolate has been redeposited at the ATCC and is available as ATCC 700927. Whole-genome libraries in M13Janus and pBluescript were prepared from genomic DNA as described for genome segments used in the K-12 genome project²⁶. Random clones were sequenced using dye-terminator chemistry and data were collected on ABI377 and 3700 automated sequencers. Sequence data were assembled by Seqman II (DNASTAR). Finishing used sequencing of opposite ends of linking clones, several PCR-based techniques and primer walking. Whole-genome optical maps for restriction enzymes *NheI* and *XhoI* were prepared²⁷ so that the ordering of contigs during assembly could be confirmed. Two gaps remain in the genome sequence. Extended exact matches pose a significant assembly challenge. The final determination of sequence for the 100-kb duplicated region was based on clones that span the junction between unique flanking sequences and the ends of the duplicated island, concordance of the two regions in optical restriction maps, excess random sequence coverage in the duplicated region, lack of polymorphism and confirmation of duplication of an internal segment by Southern blotting (data not shown).

Sequence features and database searches

Potential open reading frames (ORFs) were defined by GeneMark.hmm²⁸. The GenPept118 protein and MG1655 protein and DNA databases were searched by each ORF using BLAST²⁹. Annotations were created from the search output in which each gene was inspected, assigned a unique identifier, and its product classified by functional group⁵. Alternative start sites were chosen to conform to the annotated MG1655 sequence. Orthology was inferred when matches for EDL933 genes in the MG1655 database exceeded 90% nucleotide identity, alignments included at least 90% of both genes, and the MG1655 gene did not have an equivalent match elsewhere in the EDL933 genome. This list

was supplemented by manual inspection of the protein-level matches in the complete GenPept database to include genes with lower similarities if they occurred within co-linear regions of the genomes. The genome sequence was compared with that of MG1655 by the maximal exact match (MEM) alignment utility, (B.M., manuscript in preparation) an adaptation of MUMmer³⁰. This program was based on suffix arrays rather than suffix trees, and exact rather than unique matches, coupled with a custom anchored-alignment algorithm that extends sequence homology into the regions separating contiguous co-linear exact matches. Inferences on biases in polymorphism patterns are based on χ^2 goodness-of-fit tests of a nested sequence of multinomial log-linear models. These predict symmetric elevated levels of A $\leftrightarrow$ G, T $\leftrightarrow$ C and G $\leftrightarrow$ T polymorphisms, above a quasi-independent baseline generated from marginal frequencies in the co-occurrence matrix of synonymous third-codon differences. Further information may be found at our Website <http://www.genome.wisc.edu/>, including a Genome Browser displaying a comparative map of EDL933 and K-12.

Received 24 July; accepted 6 November 2000.

- Riley, L. W. *et al.* Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *N. Engl. J. Med.* **308**, 681–685 (1983).
- Karmali, M. A., Steele, B. T., Petric, M. & Lim, C. Sporadic cases of haemolytic–uraemic syndrome associated with faecal cytotoxin and cytotoxin-producing *Escherichia coli* in stools. *Lancet* **ii**, 619–620 (1983).
- Mead, P. S. *et al.* Food-related illness and death in the United States. *Emerg. Infect. Dis.* **5**, 607–625 (1999).
- Su, C. & Brandt, L. J. *Escherichia coli* O157:H7 infection in humans. *Ann. Intern. Med.* **123**, 698–714 (1995).
- Blattner, F. R. *et al.* The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1474 (1997).
- Reid, S. D., Herbelin, C. J., Bumbaugh, A. C., Selander, R. K. & Whittam, T. S. Parallel evolution of virulence in pathogenic *Escherichia coli*. *Nature* **406**, 64–67 (2000).
- Blattner, F. R. *et al.* Comparative genome sequencing of *E. coli* O157:H7 versus *E. coli* K 12. *Microb. Compar. Genom.* **2**, 174 (1997).
- Hacker, J. *et al.* Deletions of chromosomal regions coding for fimbriae and hemolysins occur in vitro and in vivo in various extraintestinal *Escherichia coli* isolates. *Microb. Pathog.* **8**, 213–225 (1990).
- Perna, N. T. *et al.* Molecular evolution of a pathogenicity island from enterohemorrhagic *Escherichia coli* O157:H7. *Infect. Immun.* **66**, 3810–3817 (1998).
- Plunkett, G. III, Rose, D. J., Durfee, T. J. & Blattner, F. R. Sequence of Shiga toxin 2 phase 933W from *Escherichia coli* O157:H7: Shiga toxin as a phage late-gene product. *J. Bacteriol.* **181**, 1767–1778 (1999).
- Campbell, A. & Botstein, D. in *Lambda II* (eds Hendrix, R. W., Roberts, J. W., Stahl, F. W. & Weisberg, R. A.) 365–380 (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1983).
- O'Brien, A. D. *et al.* Shiga-like toxin-converting phages from *Escherichia coli* strains that cause hemorrhagic colitis or infantile diarrhea. *Science* **226**, 694–696 (1984).
- Walterspiel, J. N., Ashkenazi, S., Morrow, A. L. & Cleary, T. G. Effect of subinhibitory concentrations of antibiotics on extracellular Shiga-like toxin I. *Infection* **20**, 25–29 (1992).
- Neely, M. N. & Friedman, D. I. Functional and genetic analysis of regulatory regions of coliphage H-19B: location of shiga-like toxin and lysis genes suggest a role for phage functions in toxin release. *Mol. Microbiol.* **28**, 1255–1267 (1998).
- Grif, K., Dierich, M. P., Karch, H. & Allerberger, F. Strain-specific differences in the amount of Shiga toxin released from enterohemorrhagic *Escherichia coli* O157 following exposure to subinhibitory concentrations of antimicrobial agents. *Eur. J. Clin. Microbiol. Infect. Dis.* **17**, 761–766 (1998).
- McDonough, M. A. & Butters, J. R. Spontaneous tandem amplification and deletion of the shiga toxin operon in *Shigella dysenteriae* 1. *Mol. Microbiol.* **34**, 1058–1069 (1999).
- Lawrence, J. G. & Ochman, H. Amelioration of bacterial genomes: rates of change and exchange. *J. Mol. Evol.* **44**, 383–397 (1997).
- Francino, M. P., Chao, L., Riley, M. A. & Ochman, H. Asymmetries generated by transcription-coupled repair in enterobacterial genomes. *Science* **272**, 107–109 (1996).
- Blaisdell, J. O., Hatahet, Z. & Wallace, S. S. A novel role for *Escherichia coli* endonuclease VIII in prevention of spontaneous G $\rightarrow$ T transversions. *J. Bacteriol.* **181**, 6396–6402 (1999).
- Le Page, F. *et al.* Transcription-coupled repair of 8-oxoguanine: requirement for XPG, TFIIH, and CSB and implications for Cockayne syndrome. *Cell* **101**, 159–171 (2000).
- Boyd, E. F., Li, J., Ochman, H. & Selander, R. K. Comparative genetics of the *inv-spa* invasion gene complex of *Salmonella enterica*. *J. Bacteriol.* **179**, 1985–1991 (1997).
- Sharp, P. M., Kelleher, J. E., Daniel, A. S., Cowan, G. M. & Murray, N. E. Roles of selection and recombination in the evolution of type I restriction-modification systems in enterobacteria. *Proc. Natl Acad. Sci. USA* **89**, 9836–9840 (1992).
- Alm, R. A. & Trust, T. J. Analysis of the genetic diversity of *Helicobacter pylori*: the tale of two genomes. *J. Mol. Med.* **77**, 834–846 (1999).
- Read, T. D. *et al.* Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39. *Nucleic Acids Res.* **28**, 1397–1406 (2000).
- Ohnishi, M. *et al.* Chromosome of the enterohemorrhagic *Escherichia coli* O157:H7: comparative analysis with K-12 MG1655 revealed the acquisition of a large amount of foreign DNAs. *DNA Res.* **6**, 361–368 (1999).
- Mahillon, J. *et al.* Subdivision of *Escherichia coli* K-12 genome for sequencing: manipulation and DNA sequence of transposable elements introducing unique restriction sites. *Gene* **223**, 47–54 (1998).
- Lin, J. *et al.* Whole-genome shotgun optical mapping of *Deinococcus radiodurans*. *Science* **285**, 1558–1562 (1999).
- Lukashin, A. V. & Borodovsky, M. GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res.* **26**, 1107–1115 (1998).
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
- Delcher, A. L. *et al.* Alignment of whole genomes. *Nucleic Acids Res.* **27**, 2369–2376 (1999).

Acknowledgements

We thank T. Forsythe, M. Goeden, H. Kijenski, B. Leininger, J. McHugh, B. Peterson,

G. Peyrot, D. Sands, P. Soni, E. Travanty and other members of the University of Wisconsin genomics team for their expert technical assistance. This work was funded by grants from the NIH (NIAID and NCHGR), the University of Wisconsin Graduate School and the RMHC to F.R.B., the NIH (NCHGR, NIAID) to D.C.S., HHMI/OTKA to G.P., an Alfred P. Sloan/DOE Fellowship to B.M., a CDC/APHL Fellowship to P.S.E., and an Alfred P. Sloan/NSF Fellowship to N.T.P. Sixteen University of Wisconsin undergraduates participated in this work and particular thanks are due to A. Byrnes for web-site development to complement this project, and to A. Darling for programming.

Correspondence and requests for materials should be addressed to N.T.P. (e-mail: perna@ahabs.wisc.edu). The GenBank accession number for the annotated sequence is AE00517H.

Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF

Vishwanath R. Iyer^{*†‡}, Christine E. Horak^{§‡}, Charles S. Scafe^{||}, David Botstein^{||}, Michael Snyder[§] & Patrick O. Brown^{*#}

^{*} Department of Biochemistry and # Howard Hughes Medical Institute, Stanford University Medical Center, Stanford, California 94305, USA

^{||} Department of Genetics, Stanford University Medical Center, Stanford, California 94305, USA

[§] Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, Connecticut 06520, USA

[†] These authors contributed equally to this work

Proteins interact with genomic DNA to bring the genome to life; and these interactions also define many functional features of the genome. SBF and MBF are sequence-specific transcription factors that activate gene expression during the G1/S transition of the cell cycle in yeast^{1,2}. SBF is a heterodimer of Swi4 and Swi6, and MBF is a heterodimer of Mbp1 and Swi6 (refs 1, 3). The related Swi4 and Mbp1 proteins are the DNA-binding components of the respective factors, and Swi6 may have a regulatory function^{4,5}. A small number of SBF and MBF target genes have been identified^{3,6–10}. Here we define the genomic binding sites of the SBF and MBF transcription factors *in vivo*, by using DNA microarrays. In addition to the previously characterized targets, we have identified about 200 new putative targets. Our results support the hypothesis that SBF activated genes are predominantly involved in budding, and in membrane and cell-wall biosynthesis, whereas DNA replication and repair are the dominant functions among MBF activated genes^{6,11}. The functional specialization of these factors may provide a mechanism for independent regulation of distinct molecular processes that normally occur in synchrony during the mitotic cell cycle.

To identify the targets of SBF and MBF, we combined chromatin immunoprecipitation and microarray hybridization (Fig. 1). Proteins were crosslinked with formaldehyde to their target sites *in vivo*. DNA that was specifically crosslinked to either of the transcription factors was purified by immunoprecipitation using an antibody against either the native protein or an epitope tag that was fused to the protein. Polymerase chain reaction (PCR) analysis of immunoprecipitated DNA confirmed the specific association of Swi4, Swi6 and Mbp1 with several known target promoters, and other target promoters that are identified here (see Supplementary Information). After reversal of the crosslinks, immunoprecipitated DNA was amplified and fluorescently labelled with the Cy5 fluoro-

[‡] Present address: Institute of Molecular and Cellular Biology, University of Texas at Austin, Austin, Texas 78712, USA.

[#] Present address: Applied Biosystems, Foster City, California 94404, USA.

phore. We also fluorescently labelled a separate control DNA sample with the Cy3 fluorophore. Genomic target loci were identified by comparative hybridization of the immunoprecipitated and control DNA probes to a DNA microarray. The ratio of the Cy5 to Cy3 fluorescence intensities measured at each DNA element in the microarray provided a measure of the extent of binding of the transcription factor to the corresponding genomic locus. As the transcription factors were expected to interact primarily with promoters, we constructed DNA microarrays that represented all the intergenic intervals, as well as all the predicted coding sequences in the yeast genome. We plotted fluorescence ratios on a representation of the yeast genome to generate a map of the distribution of these proteins on the entire genome (see Supplementary Information).

To identify the most probable targets of SBF and MBF, we ranked genomic loci according to their fluorescence ratios. We first determined the percentile rank for each array element in each immunoprecipitation experiment, and then the median percentile rank for each element across a set of experiments. The distribution of the median percentile ranks for the 11 Swi4 immunoprecipitation experiments was bimodal (Fig. 2a). Loci with median percentile ranks above 93.8 formed a group that was distinct from the bulk of the distribution. 163 loci (representing sequences upstream of 183 genes) had median percentile ranks above this threshold. The average ranks of these 163 loci were significantly higher than would be expected if the rankings in the individual immunoprecipitations were uncorrelated ($P = 8.0 \times 10^{-151}$). The genes associated with these loci therefore represent candidate SBF targets (Fig. 3). When a putative binding site fell between a pair of divergently transcribed genes, we considered both genes as potential transcriptional targets. We use 'target' to mean putative transcriptional target genes that are inferred from our immunofluorescence-microarray results.

To identify MBF targets, we considered loci with high ranks in immunoprecipitation experiments targeting Mbp1 and Swi6. The median percentile ranks from the Mbp1 plus Swi6 immunoprecipitation experiments did not show a bimodal distribution (Fig. 2b); however, loci with median percentile ranks above 94 were highly enriched for genes with the G1/S expression profile, characteristic of known MBF targets (Fig. 2c). On the basis of this analysis, and on the distribution of known MBF target genes in the ranked list, we chose a median percentile rank of 94 as a threshold for considering loci to be MBF targets. Eighty-seven loci (upstream of 98 genes) ranked above this threshold (Fig. 3). Forty-three loci ranked above the selection thresholds for both SBF and MBF, and thus seemed to represent targets of both factors (see Methods).

Three lines of evidence support the conclusion that this survey has identified *in vivo* targets of SBF and MBF. First, the putative *in vivo* binding sites we identified are highly enriched for sequences matching the defined consensus binding sites. Second, most of the genes downstream of the putative binding sites are periodically expressed during the cell cycle with a peak in G1/S, a pattern typical of the previously identified targets of these factors. Third, putative target genes with known functions are highly enriched for functions related to DNA replication, budding and the cell cycle. We have, however, probably misidentified some genes as putative targets of these factors, and failed to identify some targets.

The putative SBF and MBF targets that we identified were highly enriched for sequences matching the consensus binding sites for SBF and MBF¹². Roughly 49% of the unique intergenic fragments that were bound by Swi4 contained a motif matching the defined SBF consensus binding site (CRGAAA), whereas only about 10% of all intergenic elements contain this site ($P = 1.0 \times 10^{-30}$). Fifty-five per cent of the unique MBF intergenic targets contained the MBF consensus binding site (ACGCGN), whereas roughly twenty per cent of all intergenic fragments have such a site ($P = 1.0 \times 10^{-9}$). Compared with the genome overall, there was also a greater

tendency for the SBF and MBF target promoters to have multiple binding sites for their respective factors. The consensus site for SBF is found with roughly equal frequency in open reading frames (ORFs) and intergenic regions in the yeast genome. However, when we used DNA microarrays that contained nearly all of the predicted ORFs, and the intergenic segments of the yeast genome, to map Swi4 targets, we observed that in virtually every case in which an ORF had a high percentile rank, its upstream intergenic fragment had a similarly high rank (data not shown). Thus, despite the frequent occurrence of consensus SBF binding sequences in coding sequences, SBF binds selectively to promoters and not to coding sequences.

We used the pattern-finding programs MEME¹³ and Consensus¹⁴ to identify sequence motifs that were enriched in the high ranking (percentile rank > 98 for SBF and > 97 for MBF) target intergenic fragments. Both of these programs identified the consensus binding sites for SBF and MBF (CGCGAAAA and ACGCGN, respectively), based solely on their overabundance in the respective target loci. This suggests that these sequence motifs are indeed the primary determinants of binding of the respective factors to promoters (data not shown; and Supplementary Information). However, not every promoter bound by SBF or MBF *in vivo* contained a recognizable consensus binding site. Moreover, most of the coding sequences and many of the promoters that contain the consensus sequences show no evidence of binding to these factors *in vivo*.

SBF binding to DNA is cell-cycle regulated, occurring maximally at G1 (refs 15, 16). We saw only a modest increase in fluorescence ratios measured for SBF target loci when cells were treated with α -factor, compared with the unsynchronized or benomyl-treated cultures, perhaps because the fraction of G1 cells in unsynchronized

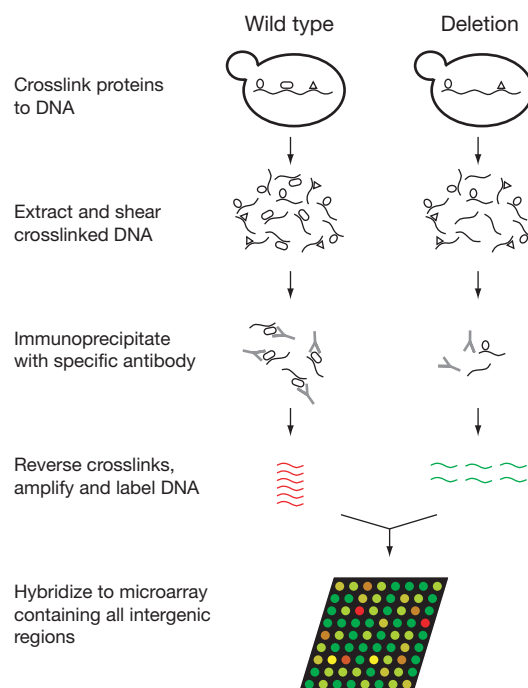


Figure 1 Strategy for analysing genome-wide protein–DNA interactions. The reference probe can either consist of DNA generated in parallel from a strain bearing a deletion of the gene encoding the protein of interest (as depicted), or of unfractionated genomic DNA amplified and labelled in the same manner. Alternatively, an epitope-tagged version of the protein of interest can be immunoprecipitated with an antibody directed against the epitope. The DNA microarray includes all of the intergenic regions or promoters from the genome. The Cy5/Cy3 fluorescence ratio for each locus reflects its enrichment by immunoprecipitation (IP) and therefore, in general, its relative occupancy by the cognate protein.

cultures or benomyl-treated cultures provided sufficient SBF-bound target sites for enrichment by immunoprecipitation.

In a comparison of the global gene expression patterns in asynchronous cultures of *swi4Δ* and wild-type cells, we found no significant differences in expression of most of the putative SBF targets, including such characterized targets as *CLN1*, *PCL1*, *PCL2* and *HO*, a result consistent with previous findings¹¹ (See Supplementary Information). Presumably, as the G1/S induction by SBF is superimposed on a basal level of SBF-independent transcription¹⁷, the effect would be more readily detected in synchronized cells. However, the small set of genes whose transcript levels were significantly reduced in the *swi4Δ* cells included several SBF targets identified by chromatin immunoprecipitation, such as *CWP1*, *CWP2*, *CIS3*, *SVS1* and *SRL1*.

We examined the expression patterns of SBF and MBF target genes during the mitotic cell cycle¹², and sporulation¹⁸, using published data (Fig. 4a). Whereas only 13% of all yeast genes are transcriptionally cell-cycle regulated¹², 66% of the SBF targets and 68% of the MBF targets, as defined by the immunoprecipitation results, were cell-cycle regulated ($P = 1.0 \times 10^{-44}$ and 1.0×10^{-22} , for SBF and MBF, respectively). Genes with maximal transcript levels during G1 and S phase showed the greatest enrichment, but there was also a less significant enrichment for genes whose transcripts peaked during other stages of the cell cycle. The putative target genes

that were not cell-cycle regulated had slightly lower immunoprecipitation-selection percentile ranks and a lower frequency of consensus MBF or SBF sites than the cell-cycle regulated targets, suggesting that at least some of these genes represent false positives in our immunoprecipitation assay. This bias was more pronounced for the putative MBF targets. There was, however, a significant enrichment for the presence of SBF consensus sites in the promoters of SBF target genes that were not cell-cycle regulated. In the set of MBF target genes that were not cell-cycle regulated, there was a significant enrichment for genes with roles in DNA repair and replication, as described below.

The known functions of many of the putative targets that we have identified raised interesting questions about the regulation of the cell cycle (Fig. 3). The G1 cyclins *CLN1* and *CLN2*, as well as *PCL1* and *PCL2*, were among the SBF targets that we expected to find. Unexpectedly, the promoters of the G2/M-specific B-type cyclins *CLB1* and *CLB2* also seemed to be bound by Swi4 in our microarray analysis, and in a confirmatory PCR analysis (data not shown). The significance of this result is unclear; SBF may be involved in repressing the transcription of these genes during G1/S or, as the mitotic B-type cyclins are required to shut off SBF-induced transcription during G2 and M phase¹⁵, this observation may point to a negative feedback loop in SBF regulation.

Why are two different transcription factors used to mediate nearly identical transcriptional programmes during the cell-division cycle in yeast? A possible answer is suggested by differences in the functions of the genes that they regulate. Many of the known transcriptional targets of SBF have roles in cell-wall biogenesis and budding¹¹. Indeed, a remarkably large fraction of the putative SBF targets identified in this study encode proteins that are involved in various aspects of bud formation, and synthesis of membrane and cell-wall components, based on functional definitions in the MIPS (<http://www.mips.biochem.mpg.de/proj/yeast/>) and YPD (<http://www.proteome.com>) databases. Twenty-seven per cent of the known genes in the SBF target set fell in these functional categories, which is a significant enrichment ($P = 1.0 \times 10^{-5}$) over their representation in the genome as a whole (11%). Twenty-five per cent of the putative MBF target genes have known roles in DNA replication, recombination and repair (on the basis of MIPS and YPD definitions), which is a significant enrichment ($P = 1.0 \times 10^{-5}$) over their representation in the genome overall (6%)². There was no enrichment of DNA metabolism-related genes among SBF targets, nor were cell-wall morphogenesis genes over-represented in the MBF targets.

The functional specialization of the putative targets of SBF and MBF was accompanied by differences in their overall expression programmes that suggest a possible regulatory logic to this division (Fig. 4b). The molecular programmes required for cellular growth and for the characteristic events at successive stages of the cell cycle occur with stereotyped synchronicity during simple mitotic proliferation. In different cell types, or during different physiological or developmental processes, however, the same molecular programmes may need to be orchestrated in different ways. In mitotically proliferating yeast, budding and DNA replication occur simultaneously; however, during the meiotic S phase, DNA replication occurs without budding. The converse is seen during mating (shmoo formation) and during pseudohyphal or invasive growth, when synthesis of membrane and cell-wall components occurs without concomitant DNA replication. Yeast cells must therefore have a mechanism to regulate genes devoted to DNA replication separately from those with major roles related to budding. As noted previously¹⁹, we observed that many of the putative MBF targets were genes that are strongly induced during sporulation (Fig. 4; data from ref. 18). The product of the B-type cyclin *CLB6*, a prominent MBF target, is also required for initiation of S phase during meiosis²⁰, and several other MBF targets involved in DNA replication and repair have crucial roles in meiosis.

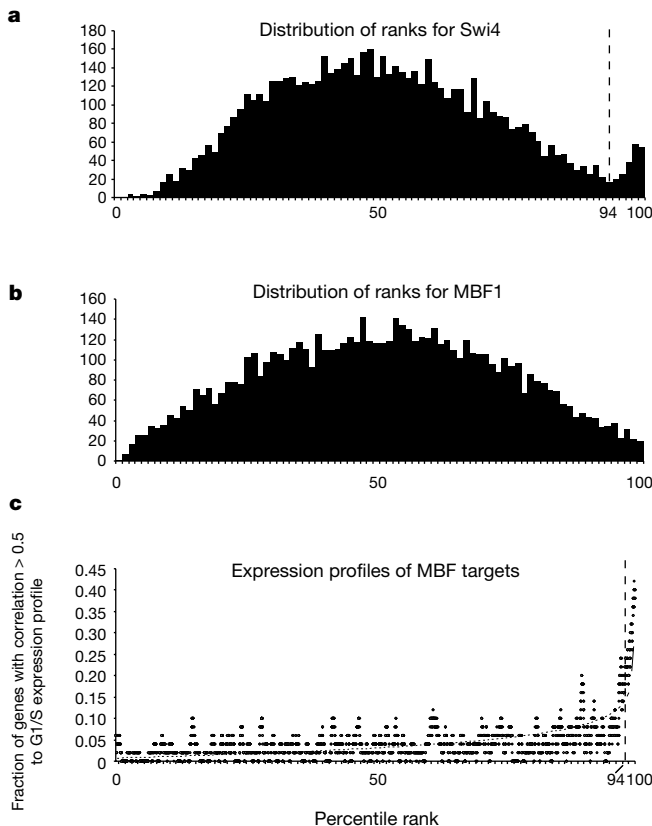


Figure 2 Defining criteria for identifying SBF and MBF targets. **a**, Distribution of median percentile ranks of Cy5/Cy3 fluorescence ratios from 11 independent microarray hybridizations analysing Swi4 binding. The threshold for defining SBF targets is indicated by a dotted line. **b**, Distribution of median ranks of Cy5/Cy3 fluorescence ratios from four independent microarray hybridizations analysing Swi6 and Mbp1 binding. **c**, Fraction of genes in a sliding 50-gene window whose expression profile has a correlation coefficient of ≥ 0.5 with that of a canonical G1/S regulated gene, as a function of the median percentile ranks of Cy5/Cy3 fluorescence ratios from the MBF hybridizations. The threshold for defining MBF targets is indicated by a dotted line.

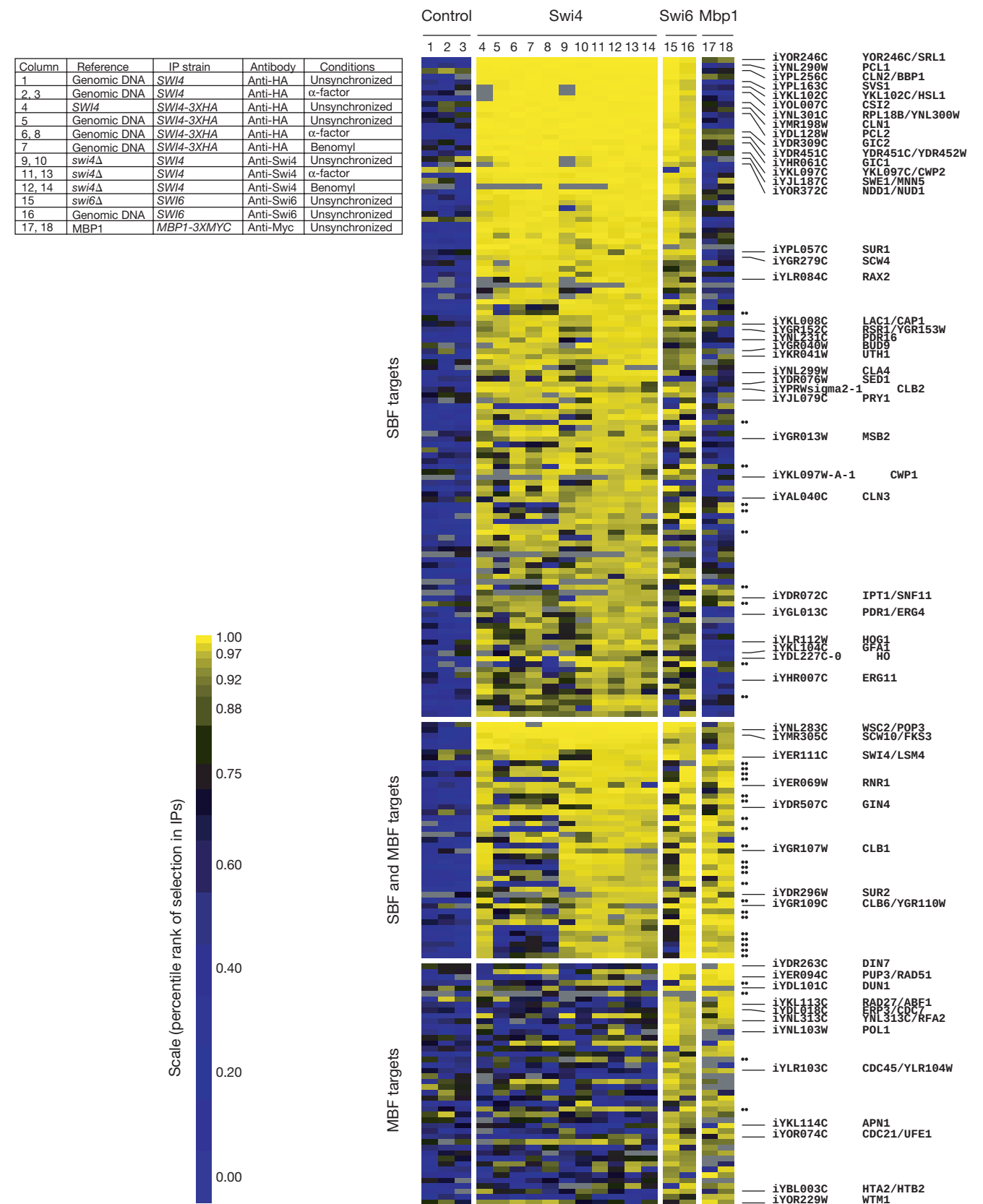


Figure 3 Genomic targets of SBF and MBF. Percentile ranks of intergenic fragments that meet selection thresholds are indicated (blue–yellow colour scale). Loci with ~70% overall nucleotide sequence identity to another yeast locus (potentially crosshybridizing) are indicated (closed circles). The combination of Cy3 and Cy5 labelled probes, the antibody used for IP (if used) and the culture conditions for each experiment are summarized (left panel). Experiments 9, 10, 17 and 18 involved independent crosslinking and IPs. DNA microarrays that included all yeast ORFs and other features, in addition to the intergenic fragments, were used for experiments 3, 8, 13 and 14.

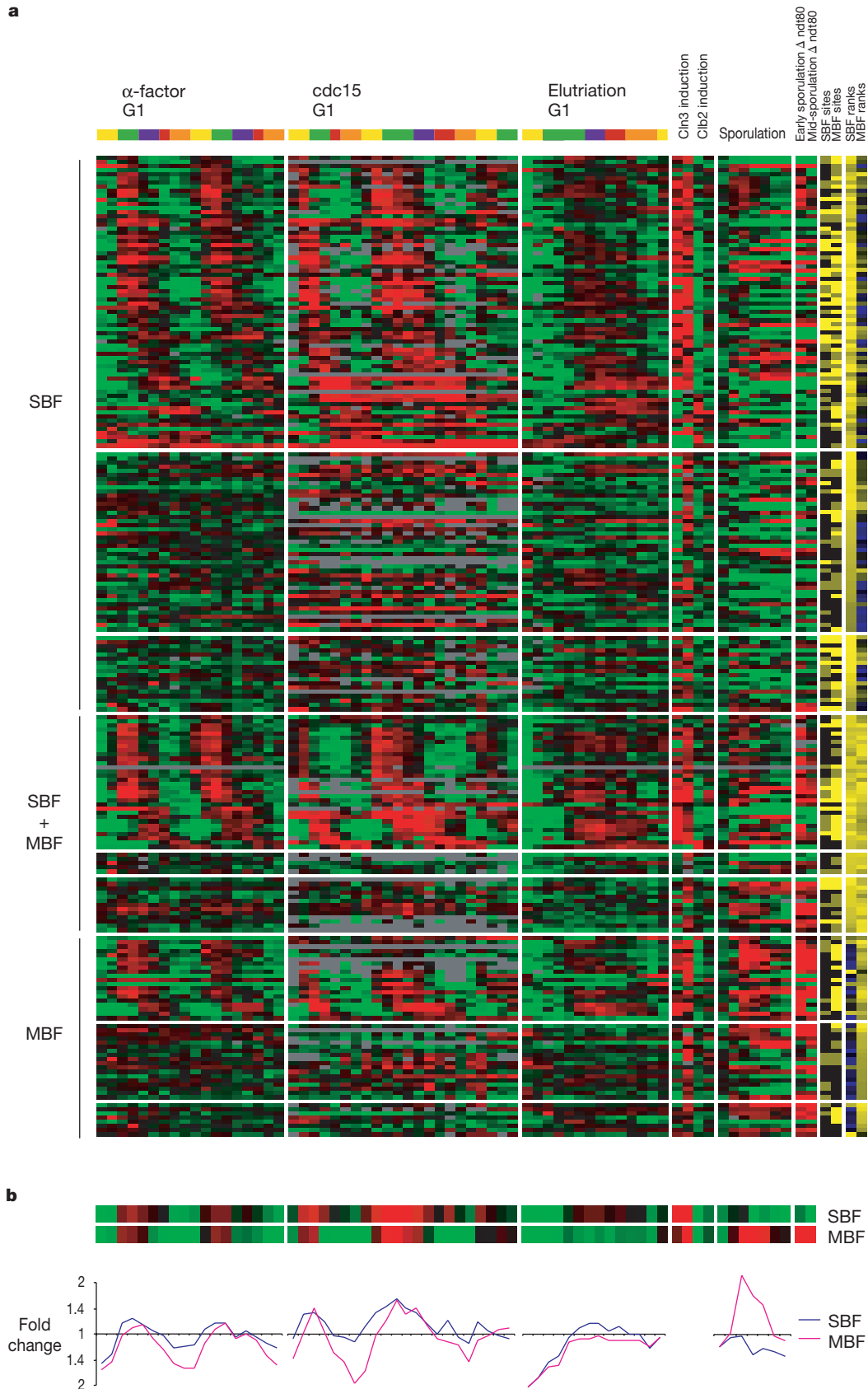


Figure 4 Expression profiles of SBF and MBF targets. **a**, Expression patterns of SBF and MBF targets are indicated (red–green colour scale). Cell-cycle data are from ref. 12 and sporulation data are from ref. 18. The stages of the cell cycle are: M/G1, yellow; G1, green; S, blue; S/G2I, red; and G2/M, orange. Yellow boxes indicate the presence of consensus binding sites in the intergenic sequences upstream of each ORF (right) and the median percentile rank in IPs of the upstream sequences is also indicated (blue–yellow

colour scale), as in Fig. 3. For each set of targets, the top panel contains cell-cycle regulated genes, the bottom panel contains genes that are members of divergently transcribed pairs in which the other member was cell-cycle regulated, and the middle panel contains the remainder of the non-cell-cycle regulated genes. **b**, Average expression profiles of the cell-cycle regulated targets of SBF and MBF, computed by averaging the $\log_2$ (Cy5/Cy3) ratios. Note the specific induction of MBF targets during sporulation.

Several of the putative SBF targets are implicated in mating and pseudohyphal growth. *RSR1* and *WSC2* have a positive role in pseudohyphal growth; *CAP1*, *GIC1* and *GIC2* are involved in shmoo formation; *SRL1*, the highest-ranking target of SBF in our immunoprecipitations, is induced by *TEC1*, a positive regulator of pseudohyphal growth; and *CLN1* is required for pseudohyphal growth²¹. The expression of *TEC1* and *PHD1*, another positive regulator of pseudohyphal growth, was reduced in the *swi4* deletion strain (see Supplementary Information), which suggests a positive function for SBF in the regulation of these genes.

Our results support a model in which SBF is the principal controller of membrane and cell-wall formation¹¹, whereas MBF primarily controls DNA replication⁶. The need for both DNA replication and membrane and cell-wall biogenesis during G1/S in dividing cells accounts for the parallel functions of SBF and MBF in the mitotic cell cycle. It is possible that SBF has an independent and distinct role in mating and pseudohyphal growth, whereas MBF has a role in meiosis. In fission yeast, an Mbp1-like factor, p73pct1, is involved in regulating both mitosis and meiosis^{22,23}. In budding yeast, it is thought that Swi6 is involved in meiosis, but to our knowledge, the role of Mbp1 in meiosis has not been examined directly²⁴. Experiments that directly examine the function of SBF and MBF in mating and meiosis, respectively, may further clarify this system.

Our results indicate that although the sequences that define canonical binding sites for SBF or MBF are common in the yeast genome, only a fraction of these sequences are actually bound by the corresponding factors *in vivo*. Moreover, most of the *in vivo* binding sites are adjacent to genes that show a characteristic cell-cycle transcription pattern. The use of DNA microarrays, together with protein-affinity isolation methods, should be generally applicable to the investigation of molecular systems that involve physical interactions of proteins with the genome. □

Methods

DNA microarrays

We synthesized roughly 6,700 PCR primer pairs to amplify nearly every intergenic region in the yeast genome. Primers were designed to amplify genomic DNA immediately adjacent to the coding sequence of every ORF or non-ORF feature, up to the end of the neighbouring ORF or non-ORF feature. Intergenic elements were designated by prefixing the SGD (Saccharomyces Genome Database, <http://genome-www.stanford.edu/saccharomyces>) designated name of the ORF or non-ORF feature that was immediately to its left with an 'i'. The primers and their sequences can be obtained from Research Genetics (Huntsville, Alabama). We prepared DNA microarrays as described²⁵.

Crosslinking and immunoprecipitation

For chromatin immunoprecipitations, 5×10^8 cells were used with the conditions as described²⁶. Briefly, cells were treated with formaldehyde to crosslink proteins and nucleic acids, and then lysed. Extracts were sonicated with a Branson 350 Sonifier with a microtip, at a power setting of 7 and a 60% duty cycle. Samples were pulsed eight times for 12 s to shear chromatin to a final average size of about 0.5–2 kilobases. We used antibodies at a 1:100 dilution for immunoprecipitation. Anti-haemagglutinin antibodies and anti-Myc antibodies were from BabCo (Berkeley Antibody Company). Affinity-purified anti-Swi6 polyclonal antibody was a gift from K. Baetz and B. Andrews. The chromatin-antibody complexes were precipitated with Protein A Sepharose beads (Pierce) and washed twice with lysis buffer and twice with 1× TBS (150 mM NaCl, 20 mM Tris-HCl, pH 7.6) for 5 min each. Chromatin was eluted from the beads with 1% SDS, 1× TE (50 mM Tris-HCl, 10 mM EDTA) and washed with 0.67% SDS, 1× TE. Crosslinks in the immunoprecipitated chromatin were reversed by heating at 65°C for at least 10 h and the DNA was purified as described²⁷, and resuspended in 10 µl of 1× TE.

Microarray hybridizations and analysis

A detailed protocol for PCR amplification and fluorescent labelling of immunoprecipitated DNA is available at <http://www.microarrays.org>. Microarrays were hybridized at 65°C for 6 h and washed as described previously, then scanned with a GenePix 4000A scanner (Axon Instruments). We quantified fluorescence intensities using GenePix Pro software (version 3.0), and uploaded data to a custom database for analysis. Data were filtered to exclude spots with obvious defects, or where the absolute signal intensity in both channels was below an empirically determined threshold. In considering loci whose percentile ranks were above the selected thresholds as putative targets, we excluded loci with similarly high ranks in control experiments (within $2 \times$ s.d. for the SBF targets, and

within $1 \times$ s.d. for the MBF targets). The control experiments were immunoprecipitates performed using antibodies specific to the HA epitope tag, carried out with a strain lacking any of the HA fusion proteins (Fig. 3, columns 1, 2 and 3).

Received 30 August; accepted 1 December 2000.

- Koch, C., Moll, T., Neuberger, M., Ahorn, H. & Nasmyth, K. A role for the transcription factors Mbp1 and Swi4 in progression from G1 to S phase. *Science* **261**, 1551–1557 (1993).
- Koch, C. & Nasmyth, K. Cell cycle regulated transcription in yeast. *Curr. Opin. Cell Biol.* **6**, 451–459 (1994).
- Andrews, B. J. & Herskowitz, I. The yeast *SWI4* protein contains a motif present in developmental regulators and is part of a complex involved in cell-cycle-dependent transcription. *Nature* **342**, 830–833 (1989).
- Primig, M., Sockanathan, S., Auer, H. & Nasmyth, K. Anatomy of a transcription factor important for the start of the cell cycle in *Saccharomyces cerevisiae*. *Nature* **358**, 593–597 (1992).
- Dirick, L., Moll, T., Auer, H. & Nasmyth, K. A central role for *SWI6* in modulating cell cycle Start-specific transcription in yeast. *Nature* **357**, 508–513 (1992).
- Lowndes, N. F., Johnson, A. L. & Johnston, L. H. Coordination of expression of DNA synthesis genes in budding yeast by a cell-cycle regulated *trans* factor. *Nature* **350**, 247–250 (1991).
- Ogas, J., Andrews, B. J. & Herskowitz, I. Transcriptional activation of *CLN1*, *CLN2*, and a putative new G1 cyclin (*HCS26*) by *SWI4*, a positive regulator of G1-specific transcription. *Cell* **66**, 1015–1026 (1991).
- McIntosh, E. M. MCB elements and the regulation of DNA replication genes in yeast. *Curr. Genet.* **24**, 185–192 (1993).
- Schwob, E. & Nasmyth, K. *CLB5* and *CLB6*, a new pair of B cyclins involved in DNA replication in *Saccharomyces cerevisiae*. *Genes Dev.* **7**, 1160–1175 (1993).
- Meadsday, V., Moore, L., Ogas, J., Tyers, M. & Andrews, B. The *PCL2* (*ORFD*)-*PHO85* cyclin-dependent kinase complex: a cell cycle regulator in yeast. *Science* **266**, 1391–1395 (1994).
- Igual, J. C., Johnson, A. L. & Johnston, L. H. Coordinated regulation of gene expression by the cell cycle transcription factor Swi4 and the protein kinase C MAP kinase pathway for yeast cell integrity. *EMBO J.* **15**, 5001–5013 (1996).
- Spellman, P. T. *et al.* Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol. Biol. Cell* **9**, 3273–3297 (1998).
- Bailey, T. L. & Elkan, C. In *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology* 28–36 (American Association for Artificial Intelligence, Menlo Park, California, 1994).
- Stormo, G. D. & Hartzell, G. W. Identifying protein-binding sites from unaligned DNA fragments. *Proc. Natl Acad. Sci. USA* **86**, 1183–1187 (1989).
- Koch, C., Schleiffer, A., Ammerer, G. & Nasmyth, K. Switching transcription on and off during the yeast cell cycle: Cln/Cdc28 kinases activate bound transcription factor SBF (Swi4/Swi6) at start, whereas Clb/Cdc28 kinases displace it from the promoter in G2. *Genes Dev.* **10**, 129–141 (1996).
- Harrington, L. A. & Andrews, B. J. Binding to the yeast Swi4,6-dependent cell cycle box, CACGAAA, is cell cycle regulated *in vivo*. *Nucleic Acids Res.* **24**, 558–565 (1996).
- Breeden, L. & Mikesell, G. Three independent forms of regulation affect expression of *HO*, *CLN1* and *CLN2* during the cell cycle of *Saccharomyces cerevisiae*. *Genetics* **138**, 1015–1024 (1994).
- Chu, S. *et al.* The transcriptional program of sporulation in budding yeast. *Science* **282**, 699–705 (1998).
- Johnston, L. H., Johnson, A. L. & Barker, D. G. The expression in meiosis of genes which are transcribed periodically in the mitotic cell cycle of budding yeast. *Exp. Cell Res.* **165**, 541–549 (1986).
- Dirick, L., Goetsch, L., Ammerer, G. & Byers, B. Regulation of meiotic S phase by Ime2 and a Clb5,6-associated kinase in *Saccharomyces cerevisiae*. *Science* **281**, 1854–1857 (1998).
- Madhani, H. D., Galitski, T., Lander, E. S. & Fink, G. R. Effectors of a developmental mitogen-activated protein kinase cascade revealed by expression signatures of signalling mutants. *Proc. Natl Acad. Sci. USA* **96**, 12530–12535 (1999).
- Zhu, Y., Takeda, T., Nasmyth, K. & Jones, N. pct1⁺, which encodes a new DNA-binding partner of p85cdc10, is required for meiosis in the fission yeast *Schizosaccharomyces pombe*. *Genes Dev.* **8**, 885–898 (1994).
- Ayté, J., Leis, J. F. & DeCaprio, J. A. The fission yeast protein p73res2 is an essential component of the mitotic MBF complex and a master regulator of meiosis. *Mol. Cell Biol.* **17**, 6246–6254 (1997).
- Leem, S. H., Chung, C. N., Sunwoo, Y. & Araki, H. Meiotic role of *SWI6* in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **26**, 3154–3158 (1998).
- DeRisi, J. L., Iyer, V. R. & Brown, P. O. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278**, 680–686 (1997).
- Strahl-Bolsinger, S., Hecht, A., Luo, K. & Grunstein, M. *SIR2* and *SIR4* interactions differ in core and extended telomeric heterochromatin in yeast. *Genes Dev.* **11**, 83–93 (1997).
- Orlando, V. & Paro, R. Mapping Polycomb-repressed domains in the bithorax complex using *in vivo* formaldehyde cross-linked chromatin. *Cell* **75**, 1187–1198 (1993).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) and the Stanford University Web site (<http://genome-www.stanford.edu/chromatinip/>) or as paper copy at the London editorial office of *Nature*.

Acknowledgements

We thank J. DeRisi for mapping software; O. Aparacio for immunoprecipitation protocols; A. Khodursky for help with significance tests; and J. Lieb for comments on the manuscript and discussion. This work was supported by grants from the National Institutes of Health, and by the Howard Hughes Medical Institute. P.O.B. is an associate investigator of the Howard Hughes Medical Institute.

Correspondence and requests for materials should be addressed to P.O.B. (e-mail: pbrown@cmgm.stanford.edu) or M.S. (e-mail: michael.snyder@yale.edu).

Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle

Holger Stark, Prakash Dube, Reinhard Lührmann & Berthold Kastner

Institut für Molekularbiologie und Tumorforschung, Emil-Mannkopfstrasse 2, 35037 Marburg, Germany
Max-Planck-Institut für biophysikalische Chemie, Am Fassberg 11, 37077 Göttingen, Germany

In eukaryotic cells, freshly synthesized messenger RNA (pre-mRNA) contains stretches of non-coding RNA that must be excised before the RNA can be translated into protein. Their removal is catalysed by the spliceosome, a large complex formed when a number of small nuclear ribonucleoprotein particles (snRNPs) bind sequentially to the pre-mRNA. The first snRNP to bind is called U1; other snRNPs (U2, U4/U6 and U5) follow¹. Here we describe the three-dimensional structure of human U1 snRNP, determined by single-particle electron cryomicroscopy at 10 Å resolution. The reconstruction reveals a doughnut-shaped central element that accommodates the seven Sm proteins common to all snRNPs, supporting a proposed model of circular Sm protein arrangement². By taking earlier biochemical results into account, we were able to assign the remaining density of the map to the other known components of U1 snRNP, deriving a structural model that describes the three-dimensional arrangement of proteins and RNA in U1 snRNP.

U1 snRNP initiates spliceosome assembly by binding to the 5' splice site through base pairs between the single-stranded 5' terminal sequence of the U1 RNA and a conserved stretch of six nucleotides at the 5' splice site of the pre-mRNA¹. This leads to functional pairing of the intronic 5' and 3' splicing sites in the early stage (commitment or E complex) of spliceosome assembly. U1 snRNP also stimulates polyadenylation of a downstream 3' end-processing³ site, is a target of the apoptotic pathway⁴, and is a potent autoantigen in systemic autoimmune diseases such as systemic lupus erythematosus⁵. Its structure is less complex than those of other snRNPs, but all spliceosomal snRNPs share common structure-determining principles⁶. Thus, the structure of the U1 snRNP is a potential model for the basic architecture of U snRNPs in general. Mammalian U1 snRNP consists of ten different proteins and a 165-nucleotide RNA molecule. Seven of the proteins, the Sm proteins (B, D1, D2, D3, E, F and G), are common to all U snRNPs, whereas the other three (U1-70K, U1-A and U1-C) are specific to U1 (Fig. 1a, b). The U1 RNA possesses the characteristic Sm site (the single-stranded sequence AAUUGUGG) to which the Sm proteins bind, generating the Sm core RNP⁷. All seven Sm proteins possess a structural Sm domain with a relative molecular mass of around 9,000 ($M_r \sim 9K$), which contains the two conserved sequence motifs Sm1 and Sm2 (reviewed in ref. 8). They also form several stable RNA-free heteromers (D1–D2, D3–B and F–E–G)⁹. The atomic structures of D1–D2 and D3–B have been determined by X-ray crystallography²; all four proteins have nearly identical Sm-domain folds, and the two dimers show the same Sm-domain interaction principle. Thus, the seven small Sm proteins appear to form a core with the snRNA, whereas the larger, particle-specific proteins determine the specific structure of the snRNP particle^{6,10}. The two largest U1-specific proteins, U1-70K and U1-A, bind directly to U1 RNA stem/loop I and II, respectively¹¹ (Fig. 1c). The structure of the amino-terminal fragment of U1-A in complex with the loop of stem II has been solved by X-ray crystallography¹². The smaller U1-C protein is probably attached by protein–protein interaction¹³.

For structural studies, we incubated functional active human U1

snRNPs with ATP to fully phosphorylate the U1-70K protein¹⁴; they were then vitrified and imaged with the electron cryomicroscope. By single-particle averaging, angular reconstitution and three-dimensional reconstruction, we determined the structure of the U1 snRNP to 10 Å resolution (Fig. 2a, b). The reconstruction presented here is the smallest ($M_r \sim 240K$) unstained three-dimensional structure that has been determined by single-particle electron cryomicroscopy, to our knowledge; it is close to the theoretical lower size limit of this technique¹⁵.

The most obvious structural feature is a ring-shaped main body (Fig. 2b) with an outer diameter of 70–80 Å. The central hole appears funnel-shaped, with its smallest diameter close to the bottom of the main body. Two large density elements (Fig. 2, labelled 70K and A) and a smaller one (IV) protrude from the top of this ring. The two large protuberances have diameters of up to 50 Å. One of them (70K) shares a large interface with the main body, whereas the other (A) connects to the ring by a thin neck only. They touch each other about 20 Å above the plane of the ring. A fourth globular element (4W) protrudes more radially from the main body and is located below protuberance A. All protruding structural elements are clustered in one area of the main body. The envelope of the reconstruction provides necessary and sufficient constraints for the interpretation of the overall architecture of the U1 snRNP by combining available structural and biochemical data (Fig. 3a–c). This was done in four steps.

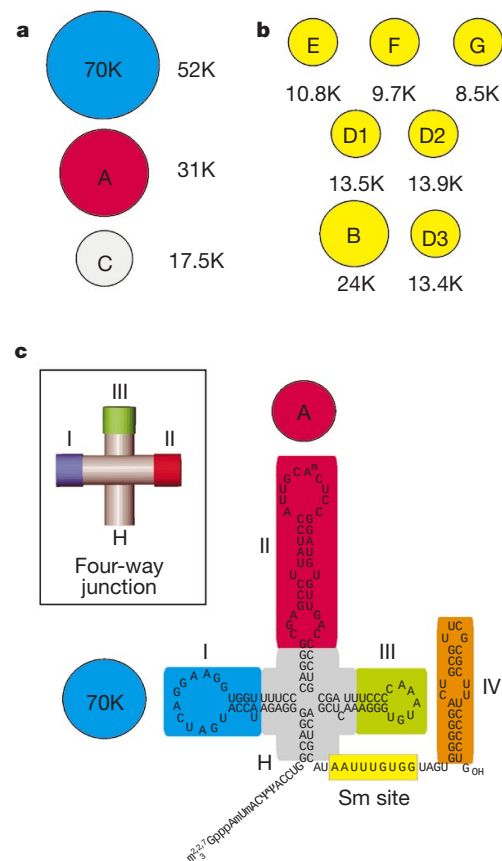


Figure 1 The RNA and protein components of U1 snRNP. **a**, The three proteins specific to U1 snRNP. **b**, The seven Sm proteins common to all snRNPs. The Sm proteins bind to the Sm-site RNA (see **c**, yellow) to generate the Sm core RNP. Numbers indicate the relative molecular masses. **c**, The secondary structure of the RNA and the sites at which proteins U1-70K and U1-A bind to the RNA. The Sm site is yellow and the four-way junction of four double-helical strands is grey. Inset, the stacking scheme of RNA double helices in the four-way junction, with stems I (blue), II (red), III (green), and the 5'–3' connecting helix H.

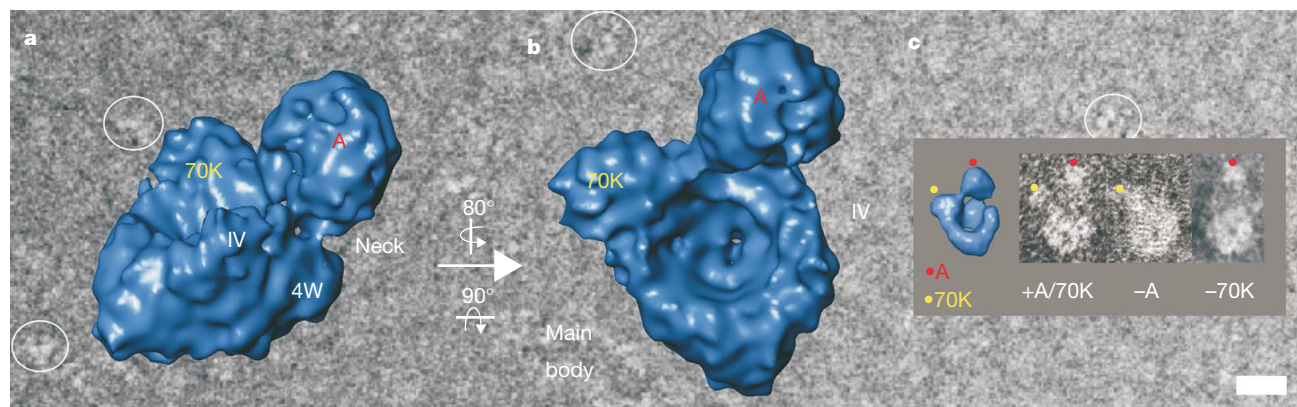


Figure 2 Three-dimensional structure of human U1 snRNP at 10 Å resolution. **a**, Side view; **b**, top view. The main body and protrusions are marked. The neck connects protrusions A and 4W. **c**, Structural discrimination between the two large protruberances. Three negative-stain electron microscopy images of U1 snRNP in approximately the same orientation and size are shown. Left, a complete U1 snRNP; centre, U1 lacking the protein U1-A; right, U1 lacking U1-70K (prepared as described^{9,16}). The protuberance created by

U1-70K is in close contact with the main body, whereas that containing U1-A connects by a narrow neck. Far left, the corresponding protuberances are marked on the lower resolution U1 snRNP reconstruction (25 Å resolution, non-ATP-incubated U1 snRNP). Background to figure, a typical electron cryomicroscopy field with some individual U1 snRNP particles circled. Scale bar, 2 nm for the U1 structure and 13 nm for the background.

First we identified the ring-shaped main body unambiguously as the Sm core RNP. The identity is strongly indicated by negative-stain electron microscopy of U1 snRNPs depleted of all specific proteins¹⁶; the structure of a depleted U1 snRNP has exactly the same size and shape as the main body of a complete snRNP. On the basis of the structures of two Sm dimers (D1–D2 and D3–B, solved by X-ray crystallography²), the interaction principle of the Sm protein interface has been extrapolated and a pseudo-symmetric heptameric ring model for the seven Sm proteins has been proposed². The modelled ring structure has an outer diameter of 70 Å and a funnel-like hole. A low-pass filtered volume representation of the model, basically containing the seven Sm-domains (provided by C. Kambach and K. Nagai), fits very well into the main body of the U1 snRNP reconstruction. The central funnel-like cavity, visible in the U1 snRNP structure as well as in the Sm-ring model, allows a precise translational alignment. Furthermore, the inherent asymmetry of the funnel reveals the up/down orientation, with the L4 loops of the Sm-domains² facing the bottom of the molecule (Fig. 3a, d). Thus, the three-dimensional image provides experimental evidence for a single-layered ring of Sm proteins in the Sm core. Most of the main body volume is filled by the heptameric Sm ring model, except for a large area in front of the ring, a nearly circular thin structure narrowing the hole and a filamentous density crossing the hole (Fig. 3d).

In the second step we located the two large U1-specific proteins, U1-A and U1-70K. Electron micrographs of negatively stained U1 snRNPs lacking either the U1-A or the U1-70K protein and antibody-labelled U1 snRNP¹⁶ have shown that the two U1-specific proteins can be assigned to the two largest protruberances of the particle. The U1-A protein could be attributed to the density that protrudes further from the main body, and the U1-70K protein corresponds to the density sharing a large interface with the Sm ring¹⁶ (Fig. 2c). This is also in line with binding and crosslinking data that demonstrate close contact between the Sm proteins and U1-70K only¹³.

The third step was to identify the structural elements of U1 RNA. On the basis of established protein–RNA interaction data¹¹, we assigned stem/loop I to the density close to U1-70K and stem/loop II to the density close to U1-A. Stems I, II and III and the 5'–3' connecting helix H of U1 RNA meet in a four-way helical junction with a defined coaxial stacking topology (Fig. 1c), as inferred from biochemical and RNA modelling studies^{17–19}. Because stems I and II of U1 RNA meet at the junction, and as their loops are in contact with U1-A and U1-70K, respectively, the only density element

assignable to the four-way junction is the one immediately below the protuberance assigned to U1-A (Fig. 3b, d; see also Fig. 2a, label 4W).

The position and shape of this density element together with the geometry of the four-way junction^{17–19} predict some bending (~45°) at the large internal loop of stem II; they cause stem/loop III to be exposed and orientated up toward the protuberances; and they make helix H point down towards the central hole of the Sm ring (Fig. 3b, e). The 3' nucleotides of helix H are buried and pack against the Sm-ring, whereas stem/loop III is surface accessible (Fig. 3b); this agrees with data from RNA structure probing^{20,21}.

The remaining 3'-terminal stem/loop IV was placed into the only remaining free density of the three-dimensional map, which had a size and shape suited to accommodate a ~40 Å double-helical RNA element (Fig. 2a, b, label IV). Stem IV was therefore fitted into the small protrusion at the top of the ring next to protein U1-A and opposite protein U1-70K (Fig. 3b). This position on the main body, clearly separated from the two large protruberances, agrees well with a previous immunoelectron microscopy localization of stem/loop IV²². The right-handed helical groove expected for a double-helical RNA element can be seen (Fig. 2b, label IV; Fig. 3d) and confirms the handedness of the reconstructed volume. The orientation of the stem IV helix shows one side in contact with the Sm ring and the other side exposed; this agrees with nuclease protection data where the 5'-terminal nucleotides of stem IV are better protected than its 3'-terminal nucleotides^{20,23}.

The positions of stem/loop IV at the top of the Sm ring and of helix H close to the bottom of the ring define the terminal ends of the Sm-site RNA sequence (Fig. 1). The short sequence AAUUU-GUGG of the U1 Sm-site RNA is sufficient to form the Sm core⁷ and should provide the essential Sm protein binding sites. In ref. 24, various Sm-site nucleotides could be crosslinked to Sm proteins, and the crosslinked amino acids were located in the L3 loops of the respective Sm-domain structure. Because the L3 loops of all Sm proteins project into the hole of the Sm-ring model², the Sm-site RNA should make multiple contacts to the Sm protein surface in this central cavity. According to the U1 snRNP model, the Sm-site sequence should enter the interior of the ring from the bottom where helix H is located, and lead to stem/loop IV at the top of the Sm ring (Fig. 3e). This means that the Sm-site RNA goes through the hole of the Sm ring as suggested². As the Sm-site RNA has multiple Sm protein contact sites at the interior ring surface²⁴ (see above), we propose that the circular thin density in the hole is the Sm-site RNA passing through the hole from bottom to top of the

Sm ring. In the Sm core, the Sm-G protein is efficiently crosslinked to the first uridine residue (AAU) of the Sm site²⁴. The Sm-G protein should consequently be located close to helix H, the starting point of the Sm site. This location of the Sm-G protein roughly defines the rotational orientation of the Sm ring (Fig. 3d). Adoption of the Sm protein succession in the heptameric ring from ref. 2 brings protein D2 close to U1-70K. This is also consistent with the observed crosslink between these two proteins¹³.

In the final stage of modelling, we assigned the remaining density, located in front of the ring, opposite the two large protuberances (Fig. 3c, d). It accounts for the expected volumes of the last two sizeable protein domains: protein U1-C and the large carboxy-terminal domain ($M_r \sim 15K$) of Sm-B. The continuity of the density at this position is supported biochemically by the crosslink between Sm-B and U1-C¹³, and it also allows Sm-B to be in close contact with U1-70K, as required by the observation of a crosslink between these two proteins¹³. The assignment is supported by the observations that U1-C is involved in the recognition of the 5' splice site and is expected to be located close to the 5'-end of U1 RNA²⁵, and that the 5'-terminal m₃G cap of U1 RNA has been observed by immuno-

electron microscopy¹⁶ in the area where U1-C and Sm-B are located in our model.

The molecular interpretation of the electron cryomicroscopy map obtained is self-consistent, in that no major density in the map or protein/RNA component remains unaccounted for (Fig. 3c). Furthermore, it agrees well with many biochemical results. The model not only gives a picture of the macromolecular structure of the U1 snRNP; it also provides some ideas about the general architecture of snRNPs. The most obvious characteristic is the central location of the heptameric Sm protein ring. Another common snRNP feature may be the intimate contact found in the model between the flanking regions of the Sm-site RNA and the Sm proteins. We suggest that the corresponding Sm-site RNA flanking regions in the various snRNPs will bind to similar Sm protein positions, thus leading to Sm-site RNAs passing through the hole of the heptameric Sm-ring. □

Methods

We isolated U1 snRNPs from HeLa nuclear extract and purified them as described¹⁶. Just before specimen preparation, U1 snRNPs were incubated at 4 °C for 2 h with 10 mM ATP. A solution containing 7.5 µg ml⁻¹ U1 snRNP was applied to perforated carbon grids and rapidly frozen in liquid ethane²⁶. Images were taken at a magnification of 66,000 and at three different defoci (-1.0, -1.34, -1.5 µm) using a Gatan cryostage in a Philips CM200 FEG operated at 160 kV. The micrographs were scanned with a rotating-drum scanner (Tango, Heidelberger Druckmaschinen) at a step size of 8 µm. About 22,000 single-particle images were extracted semi-automatically from the micrographs and used to calculate the three-dimensional map using the IMAGIC-V (Image Science GmbH, Berlin) software package²⁷. The filtered raw images were subjected to multi-reference alignment, multivariate statistical analysis²⁸ and automated classification to obtain class averages of characteristic views of the molecule with improved signal-to-noise ratio. Euler angle orientations of these class averages were determined by using the angular reconstitution approach. Three-dimensional reconstructions were then computed using a back-projection algorithm²⁹ (3 σ threshold, 9.5 Å and 0.5 threshold, 15 Å). The modelling and surface rendering of the structure were performed with the program ERNA-3D (Florian Müller, Garching Innovation GmbH) and Iris Explorer (NAG), respectively.

Received 22 August; accepted 17 November 2000.

- Moore, M. J., Query, C. C. & Sharp, P. A. in *The RNA World* (eds Gesteland, R. F. & Atkins, J. F.) 303–357 (Cold Spring Harbor Laboratory, Cold Spring Harbor, 1993).
- Kambach, C. et al. Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. *Cell* **96**, 375–387 (1999).
- Gundersen, S. I., Polycarpou-Schwarz, M. & Mattaj, J. W. U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol. Cell* **1**, 255–264 (1998).
- Casciola-Rosen, L. et al. Apoptin/CPP32 cleaves proteins that are essential for cellular repair: a fundamental principle of apoptotic death. *J. Exp. Med.* **183**, 1957–1964 (1996).
- van Venrooij, W. J. & Pruijn, G. J. M. Ribonucleoprotein complexes as autoantigens. *Curr. Opin. Immunol.* **7**, 819–824 (1995).
- Kastner, B., Bach, M. & Lührmann, R. Electron microscopy of small nuclear ribonucleoprotein (snRNP) particles U2 and U5: evidence for a common structure-determining principle in the major U snRNP family. *Proc. Natl Acad. Sci. USA* **87**, 1710–1714 (1990).
- Raker, V. A., Hartmuth, K., Kastner, B. & Lührmann, R. Spliceosomal U snRNP core assembly: Sm proteins assemble onto an Sm site RNA nonanucleotide in a specific and thermodynamically stable manner. *Mol. Cell. Biol.* **19**, 6554–6565 (1999).
- Kambach, C., Walke, S. & Nagai, K. Structure and assembly of the spliceosomal small nuclear ribonucleoprotein particles. *Curr. Opin. Struct. Biol.* **9**, 222–230 (1999).
- Raker, V. A., Plessel, G. & Lührmann, R. The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro. *EMBO J.* **15**, 2256–2269 (1996).
- Will, C. L. & Lührmann, R. Protein functions in pre-mRNA splicing. *Curr. Opin. Cell Biol.* **9**, 320–328 (1997).
- Nagai, K. & Mattaj, J. W. *RNA-Protein Interactions in the Splicing snRNPs* (eds Nagai, K. & Mattaj, J. W.) 150–177 (Oxford Univ. Press, Oxford, 1994).
- Oubridge, C., Ito, N., Evans, P. R., Teo, C. H. & Nagai, K. Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* **372**, 432–438 (1994).
- Neilsen, R. L., Will, C. L., van Venrooij, W. J. & Lührmann, R. The association of the U1-specific 70K and C proteins with U1 snRNPs is mediated in part by common U snRNP proteins. *EMBO J.* **13**, 4113–4125 (1994).
- Tazi, J. et al. Thiophosphorylation of U1-70K protein inhibits pre-mRNA splicing. *Nature* **363**, 283–286 (1993).
- Henderson, R. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Q. Rev. Biophys.* **28**, 171–193 (1995).
- Kastner, B., Kornstädt, U., Bach, M. & Lührmann, R. Structure of the small nuclear RNP particle U1: identification of the two structural protuberances with RNP-antigens A and 70K. *J. Cell Biol.* **116**, 839–849 (1992).
- Krol, A. et al. Solution structure of human U1 snRNA. Derivation of a possible three-dimensional model. *Nucleic Acids Res.* **18**, 3803–3811 (1990).
- Duckett, D. R., Murchie, A. I. & Lilley, D. M. The global folding of four-way helical junctions in RNA, including that in U1 snRNA. *Cell* **83**, 1027–1036 (1995).

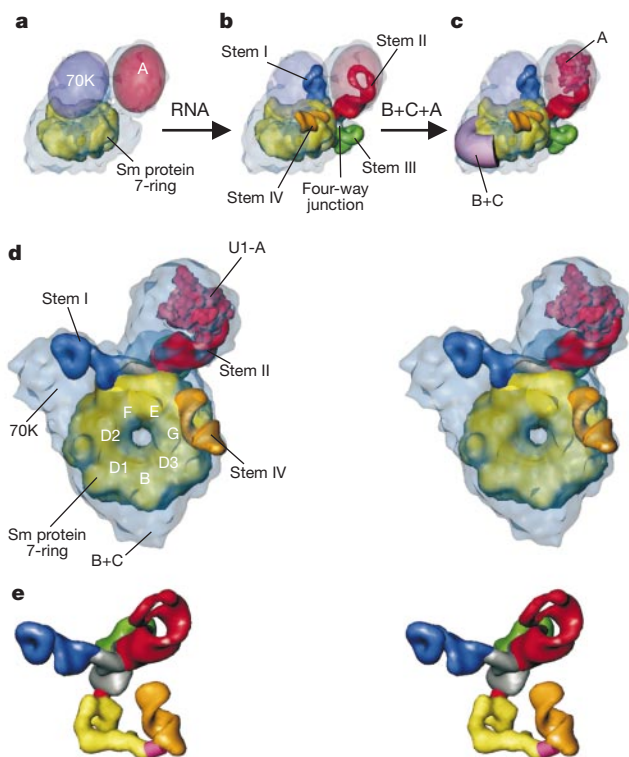


Figure 3 Modelling strategy: integration of available biochemical and structural information to fit the components of U1 snRNPs (Figs 1, 2a, b) into the three-dimensional electron cryomicroscopy maps. **a**, The U1 snRNP structure with the two main protuberances (U1-A and U1-70K, Fig. 2; spheres correspond to their expected volumes). A surface representation of the heptamer Sm ring² was fitted into the structure with translational alignment based on the shapes of the structures. **b**, RNA was modelled according to the known binding interactions between U1 RNA and U1-A and U1-70K (Fig. 1c). The four-way junction and stem III were fitted to the protrusion below U1-A. Stem IV is at the top of the Sm ring, opposite 70K. Colours as in Fig. 1. **c**, The extra domain of U1-A and U1-C were located in the map as one toroidal volume (purple), corresponding approximately to the sum of their molecular masses. The fragment of U1-A (red) was modelled according to the X-ray crystal structure of the complex¹². **d**, Enlarged stereo view of the model (viewing direction as in Fig. 2b). The rotational orientation of the Sm ring has been determined by crosslinking data. The positions of the Sm proteins are indicated as proposed². **e**, Stereo view of the modelled U1 snRNA in almost the same orientation as **d**. Helix H (grey) and stem IV (orange) are connected by the Sm-site RNA (yellow). Light red, 5' end of Sm-site RNA; pink, 3' end. The Sm site passes through the hole from bottom to top and was tentatively modelled in a spiralling way.

19. Walter, F., Murchie, A. I., Duckett, D. R. & Lilley, D. M. Global structure of four-way RNA junctions studied using fluorescence resonance energy transfer. *RNA* **4**, 719–728 (1998).
20. Liautaud, J. -P., Sri-Widada, J., Brunel, C. & Jeanteur, P. Structural organisation of ribonucleoproteins containing small nuclear RNAs from HeLa cells. *J. Mol. Biol.* **162**, 623–643 (1982).
21. Hartmuth, K., Raker, V. A., Huber, J., Branlant, C. & Lührmann, R. An unusual chemical reactivity of Sm site adenosines strongly correlates with proper assembly of core U snRNP particles. *J. Mol. Biol.* **285**, 133–147 (1999).
22. Hoet, R. M., Kastner, B., Lührmann, R. & van Venrooi, W. J. Purification and characterization of human autoantibodies directed to specific regions on U1RNP; recognition of native U1RNP complexes. *Nucleic Acids Res.* **21**, 5130–5136 (1993).
23. Lin, W. -L. & Pederson, T. Ribonucleoprotein organization of eukaryotic RNA. XXXI. Structure of the U1 small nuclear ribonucleoprotein. *J. Mol. Biol.* **180**, 947–960 (1984).
24. Urlaub, H., Raker, V., Kostka, S. & Lührmann, R. Sm protein–Sm site RNA interactions within the inner ring of the spliceosomal snRNP core structure. *EMBO J.* **20**, 1–10 (2001).
25. Heinrichs, V., Bach, M., Winkelmann, G. & Lührmann, R. U1-specific protein C needed for efficient complex formation of U1 snRNP with a 5' splice site. *Science* **247**, 69–72 (1990).
26. Adrian, M., Dubochet, J., Lepault, J. & McDowell, A. W. Cryo-electron microscopy of viruses. *Nature* **308**, 32–36 (1984).
27. van Heel, M., Harauz, G. & Orlova, E. V. A new generation of the IMAGIC image processing system. *J. Struct. Biol.* **116**, 17–24 (1996).
28. van Heel, M. & Frank, J. Use of multivariate statistics in analysing the images of biological macromolecules. *Ultramicroscopy* **6**, 187–194 (1981).
29. van Heel, M. & Harauz, G. Resolution criteria for three-dimensional reconstruction. *Optik* **73**, 119–122 (1986).

Acknowledgements

We thank F. Müller for the modelling software ERNA-3D; K. Nagai and C. Kambach for the Sm protein ring model; and M. Golas and B. Sander for assistance in electron microscopy. This work was supported by the Gottfried Wilhelm Leibniz Program and a grant from the Deutsche Forschungsgemeinschaft.

Correspondence and requests for materials should be addressed to H.S. (e-mail: Holger.Stark@mpibpc.mpg.de).

addendum

A one-hit model of cell death in inherited neuronal degenerations

G. Clarke, R. A. Collins, B. R. Leavitt, D. F. Andrews, M. R. Hayden, C. J. Lumsden & R. R. McInnes

Nature **406**, 195–199 (2000).

It has been questioned whether our neurodegeneration model was anticipated by Calne and co-workers^{1–4}. Both groups related exponential neuronal loss to pathogenesis, but the models are radically different. We recognized that in neurodegeneration, exponential death kinetics mean that the risk of cell death remains constant with age. To accommodate constant risk, we proposed a one-hit model: the hit is a *consequence* of a neuron being in a high-risk (for example, mutant) state; a hit is a biochemical reaction within each *single* neuron, committing it to death at a random time. We noted that constant risk excludes cumulative intracellular damage as a disease mechanism.

Calne *et al.* proposed that in idiopathic Parkinsonism “An event [for example, infection] acts on the neuronal *population* over a brief period ... [to initiate disease]”². They excluded ongoing processes as causing idiopathic Parkinsonism^{2,4}, but the role of cumulative intracellular damage was not addressed. Their focus on whole neuron populations and on events that act on them is notable, and complementary to but different from our model. □

1. Lee, C. S. *et al.* Clinical observations on the rate of progression of idiopathic parkinsonism. *Brain* **117**, 501–507 (1994).
2. Schulzer, M., Lee, C. S., Mak, E. K., Vingerhoets, F. J. G. & Calne, D. B. A mathematical model of pathogenesis in idiopathic parkinsonism. *Brain* **117**, 509–516 (1994).
3. Calne, D. B. Is idiopathic parkinsonism the consequence of an event or a process? *Neurology* **44**, 5–10 (1994).
4. Lee, C. S. *et al.* Patterns of asymmetry do not change over the course of idiopathic parkinsonism: implications for pathogenesis. *Neurology* **45**, 435–439 (1995).

corrections

Kainate receptors are involved in synaptic plasticity

Zuner A. Bortolotto, Vernon R. J. Clarke, Caroline M. Delany, Michael C. Parry, Ilse Smolders, Michel Vignes, Ken H. Ho, Peter Miu, Bradford T. Brinton, Robert Fantiske, Ann Ogden, Mary Gates, Paul L. Ornstein, David Lodge, David Bleakman & Graham L. Collingridge

Nature **402**, 297–301 (1999).

In this paper, the authors omitted the reference to the published synthetic pathway¹ by which the selective GluR5 kainate receptor antagonist LY382884 may be made. □

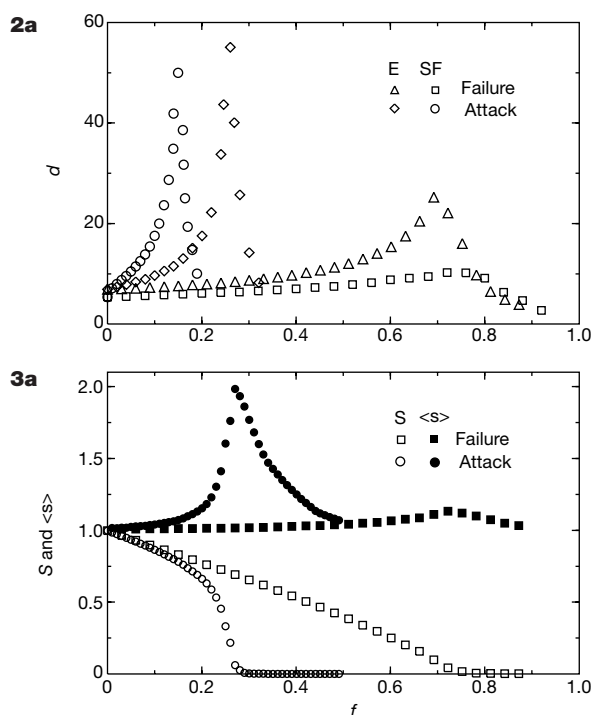
1. Bleisch, T. J. *et al.* Structure activity studies of aryl-spaced decahydroisoquinoline-3-carboxylic acid AMPA receptor antagonists. *Bioorg. Med. Chem. Lett.* **7**, 1161–1166 (1997).

Error and attack tolerance of complex networks

Réka Albert, Hawoong Jeong & Albert-László Barabási

Nature **406**, 378–382 (2000).

In this paper, the error tolerance curves for the exponential network were affected by a software error. This did not impact the attack curves nor the measurements and conclusions regarding the error/attack tolerance of scale-free networks, the World-Wide Web and the Internet. The corrected Figs 2a and 3a are shown below. □



From bioreactors to HPLC

A new take on the Dewar flask and DNA analysis columns.



Adagio from Pierre Guerin: its lips are sealed.

Adagio bioreactors

Pierre Guerin

Autoclavable bioreactors

This range of economical bioreactors includes models with agitation by turbines for bacterial studies or by an HTPG propeller for cell cultures. Air-tightness is achieved by a lip seal, or, as an option, a magnetic coupling. The vessel temperature is regulated by a thermostat-controlled water circuit in the double shell. A filtered-air intake and rotameter-regulated air supply provide aeration. Working volumes of two, five and eight litres are available. The system uses Windows-based software.

Reader Service No. 100

APS 540

Intergen

www.intergen.com

A reagent described in glowing terms

An extremely high signal-to-noise ratio resulting in a 10-fold increase in sensitivity

over current agents is claimed for this alkaline phosphatase chemiluminescence substrate. It offers a long, steady glow without signal decay, allowing the user to reach sensitivity levels of radioactivity without the associated hazards.

Reader Service No. 101

Eastjtec PLUS

Flowgen

www.flowgen.co.uk

Alarming electroporation

This unit can be used for electroporation of all cell types, including mammalian, plant, bacterial and yeast cells, and can be operated in single, double-pulse or 'optipulse' mode. A benchtop remote control unit with an LCD and 'smart card' reader/recorder for storing and reproducing programs allows experimental development in constantly changing conditions. A printer is supplied with the unit. Safety features include pre-arc detection, open and short-circuit detection, pre-

and post-pulse impedance measurements, and visual and audible alarms to warn users when incorrect or unsafe information is programmed or measured.

Reader Service No. 102

Klerwipe CR

Shield Medicare

www.shieldmedicare.com

Sterile IMS tub wipes

Made from low-lint, hydro-entangled cellulose and impregnated with 70% denatured ethanol, these sterile wipes contain no chemical binders, are not affected by commonly used chemicals or solvents, and maintain their wetness for two years. The wipes are 200 × 230 mm in size and are supplied in tubs of 200.

Reader Service No. 103

High-temperature furnaces

Lenton

www.lentonfurnaces.com

Heating power at the bottom of the range

Maximum operating temperatures of 1,700°C and 1,800°C can be reached by these bottom-loading furnaces, which use an electrically powered hearth to provide safe working conditions and facilitate handling of heavy or delicate loads. Fast heating and cooling cycles can be achieved by moving samples in and out of the chamber on the hearth as required. The furnaces can be used with atmospheres other than air, and rotating hearths are also available. Three sizes are available, ranging from 3.4 to 21 litres.

Reader Service No. 104

TOPO Walker kit

Invitrogen

www.invitrogen.com

Mind the gap

For researchers sequencing artificial chromosomal (BACs, PACs and YACs), this kit will simplify and hasten sequencing of the gaps that develop during 'shotgun' subcloning methods, eliminating library screening and probe regeneration. The kit uses a 58-bp topoisomerase-activated double-stranded linker which, when ligated to the end of an unknown DNA sequence, enables the DNA to be specifically amplified and directly sequenced, or cloned for further analysis.

Reader Service No. 105

Nalgen Dewar flasks

Nalge Nunc

www.nalgenunc.com/dewar

Get to grips with these

Nalge Nunc's new Dewar flasks have a finger-grip on the bottom, ribs for easier handling, a low centre of gravity to reduce spills, a large



Dewar like these?

opening to facilitate access and a built-in spout. The flasks and covers are made of high-density polyethylene and have a bail-type steel handle. They are available in capacities of 1, 2, 5 and 10 litres.

Reader Service No. 106

Help on the web

Bender MedSystems

www.bendermedsystems.com

Cell biologists may get stuck on this site

Intended to aid researchers working in fields such as cell adhesion, apoptosis, cytokines and tumour biology, this website includes comprehensive product information as well as application guides, crossreactivity tables and technical literature on a wide variety of antibodies, bioreagents and ready-to-use assay kits. Product listings are organized into functional categories including adhesion molecules, apoptosis research, CD markers, tumour biology, annexin V proteins, cytokines and cytokine receptors. Each of these is further subdivided into specific product types. The site also contains a product index and an interactive search facility.

Reader Service No. 107

Gene Expression DataMart

NetGenics

www.netgenics.com

A matter of integrity

DataMart can be used to clean, normalize and integrate gene expression data from any combination of sources, including oligo-array, spotted-array, RT-PCR, Taq-Man and others. An 'extraction-transformation-load' paradigm is used to deposit expression data into a single repository where it is optimized for browsing and easy access, allowing scientists to ask broad scientific questions across various types of expression data. By avoiding the need to merge multiple queries and query results, this facility offers rapid access to clean data unhindered by errors in identification, classifi-

cation, spelling and vocabulary conformation.
Reader Service No. 108

ReadyStrip

Bio-Rad Laboratories

www.bio-rad.com

A new dimension in IPG strips

These strips for first-dimension IEF separations are available in 7-, 11- and 18-cm gel lengths that will also fit all second-dimension gel systems. Choose from broad pH ranges of 3–10 or 4–7 for maximum resolution of total protein. The narrow, overlapping pH gradients (3–6, 5–8 and 7–10) allow greater resolution with more centimetres of gel per pH unit, for stand-alone use or for the creation of overlapping 'virtual' cells. Anode location and pH range are permanently labelled on each strip.

Reader Service No. 109

GeneLIMO

CPG

Versatile transfection reagents

These reagents are liposomal formulations of polycationic lipids and a neutral, non-transfecting lipid compound. GeneLIMO-Plus is used to transfect DNA into a wide variety of primary cells. It is also effective in transfection of cultured cells. GeneLIMO-Super has been successfully used to transfect DNA into CHO, COS-1 and NIH 3T3 in the presence of fetal bovine serum, and can also be employed to transfect primary cells. The reagents are formulated for both transient and stable transfections with or without the presence of serum.

Reader Service No. 110

Superscript

Life Technologies

www.lifetech.com

One-step RC-PCR

Suitable for screening of gene expression, detection of rare and viral RNA, and quantification of specific messages, these RT-PCR systems allow cDNA synthesis and amplification in a single tube. The Taq DNA polymerase sys-

tem can detect as few as 5–10 molecules of RNA template from as little as 0.1 pg, enabling amplification of rare targets or from limited starting material. A version for long templates offers the same level of sensitivity for targets up to 9kb.

Reader Service No. 111

BESS-T & G

Epicentre

www.epicentre.com/bess

Nothing special

Using this kit, researchers should be able to discover and score single nucleotide polymorphisms (SNPs) and other changes including insertions, deletions, frameshifts and repeat expansions in a PCR-amplified DNA template in as little as four hours. Detection and localisation of mutations are obtained simultaneously using a two-step procedure that requires no special gels, gel electrophoresis conditions, hazardous reagents, reannealing steps or extensive reaction optimization. The technology can be scaled up for high-throughput analysis on multichannel DNA sequencers.

Reader Service No. 112

ProStar Helix HPLC column

Varian

www.varianinc.com

Speedy separations

Promising ultra-fast separations of DNA fragments by size, the manufacturers of this high-performance liquid chromatography (HPLC) column describe it as a lower-cost alternative to gel electrophoresis that is suitable for denaturing HPLC and for the analysis and purification of PCR products and restriction digests. Researchers can expect 4,000 injections per column and a 50% decrease in buffer consumption. The ProStar can also be used for automated detection of single nucleotide polymorphisms.

Reader Service No. 113

These notes are compiled in the Nature office from information provided by the manufacturers.